

***WASHINGTON ALTERNATE ASSESSMENT SYSTEM
TECHNICAL REPORT
2006***

*Prepared for the
OFFICE OF THE SUPERINTENDENT OF PUBLIC INSTRUCTION*

*by
PEARSON EDUCATIONAL MEASUREMENT*



October 2006

Table of Contents

<i>I. EXECUTIVE SUMMARY</i> -----	4
<i>II. What is WAAS?</i> -----	5
Program Purpose-----	6
Target Population & Participation Rates-----	7
<i>III. PORTFOLIO DEVELOPMENT</i> -----	8
Construct Validity-----	9
Content Validity-----	17
<i>IV. CONTENT WORKGROUPS & DEVELOPMENT OF EALR EXTENSIONS</i> -----	17
Content Workgroups-----	17
Membership-----	18
Application and Implementaion-----	19
<i>V. SUMMER SCORING INSTITUTE</i> -----	19
<i>VI. TECHNICAL QUALITIES OF SCORING</i> -----	23
Inter-Rater Agreement-----	23
Coefficient Alpha-----	30
<i>VII. SUMMARY OF STUDENT PERFORMANCE</i> -----	32
<i>VIII. PERFORMANCE ON THE WAAS RELATIVE TO STANDARDS</i> -----	36
<i>APPENDIX A. SPECIAL EDUCATION ASSESSMENT ADVISORY PANEL—2006 MEMBERSHIP</i> -----	40
<i>APPENDIX B. SPECIAL EDUCATION ASSESSMENT LEADERSHIP TEAM (SEALT)—2006 MEMBERSHIP</i> -----	42
<i>APPENDIX C. SPECIAL EDUCATION CONTENT WORKGROUPS—2006 MEMBERSHIP</i> -----	44
<i>APPENDIX D. PARTICIPATION GUIDELINES FOR THE 2006 WAAS</i> -----	46
<i>APPENDIX E. WAAS PORTFOLIO SCORING 2006 PROCEDURAL STEPS</i> -----	48
<i>APPENDIX F. WAAS PORTFOLIO SCORING SUMMARY SHEET—2006</i> -----	50
Part I: Progress on IEP Skills (Progress on targeted skills scored separately in each content area)---	49
Part II: Student Generalization of Skills (These dimensions are scored across portfolio)-----	50
Dimension-----	50
<i>APPENDIX G. WAAS PORTFOLIO ACADEMIC ACHIEVEMENT STANDARD DESCRIPTIONS</i> -----	51

List of Tables and Figures

Table 2.1 Number of Participants in the WAAS from 2002.....	7
Table 3.1 Grade 3 Sub-score Correlations	10
Table 3.2 Grade 4 Sub-score Correlations	10
Table 3.3 Grade 5 Sub-score Correlations	11
Table 3.4 Grade 6 Sub-score Correlations	11
Table 3.5 Grade 7 Sub-score Correlations	12
Table 3.6 Grade 8 Sub-score Correlations	12
Table 3.7 Grade 10 Sub-score Correlations	13
Table 3.8 Factor Loadings	14
Table 5.1 Contents and Dimensions of the 2006 WAAS Scoring.....	20
Table 6.1 Percentage Agreement between First and Second Scorers for Grade 3	25
Table 6.2 Percentage Agreement between First and Second Scorers for Grade 4	25
Table 6.3 Percentage Agreement between First and Second Scorers for Grade 5	25
Table 6.4 Percentage Agreement between First and Second Scorers for Grade 6	26
Table 6.5 Percentage Agreement between First and Second Scorers for Grade 7	26
Table 6.6 Percentage Agreement between First and Second Scorers for Grade 8	26
Table 6.7 Percentage Agreement between First and Second Scorers for Grade 10	27
Table 6.8 Overall Percentage Agreement between First and Second Scorers	27
Table 6.9 Frequency Distribution of the Number of Scores Reached at Least Adjacent Agreement on Validity Portfolios	28
Table 6.10 Extent of Agreement between Scorers and Experts on Validity Portfolios: Grade 5	29
Table 6.11 Extent of Agreement between Scorers and Experts on Validity Portfolios: Grade 7	29
Table 6.12 Extent of Agreement between Scorers and Experts on Validity Portfolios: Grade 10	29
Table 6.13 Coefficient Alpha for Total Scores and Part II Scores	31
Table 7.1 2006 WAAS Score Distributions	33
Table 7.1 2006 WAAS Score Distributions (Continued)	34
Table 7.1 2006 WAAS Score Distributions (Continued)	35
Table 8.1 Decision Rule for Determining Level of Performance on WAAS Portfolio (Grades 3-8)	37
Table 8.2 Decision Rule for Determining Level of Performance on WAAS Portfolio (Grade 10)	37
Table 8.3 Percentage of Students by Performance Level on 2006 WAAS.....	38
Figure 8.1 2006 WAAS Performance Distribution by Gender	39

I. EXECUTIVE SUMMARY

The Washington Alternate Assessment System (WAAS) was administered operationally for the sixth year in spring 2006. In compliance with professional standards that test developers produce a technical manual to document the technical quality of an assessment and evidence for the reliability and validity of test scores (*Standards for Educational and Psychological Testing*, AERA/APA/NCME, 1999), this manual summarizes the technical information for the 2006 WAAS.

The 2006 WAAS test administration window began on April 17, 2006, and ended on May 5, 2006. All application materials were sent to the Pearson Educational Measurement (PEM) Scoring Center by May 19. All received portfolios were scored during the Summer Scoring Institute, which convened from June 19 to June 29. Scored portfolios were returned to school districts in August. With Grade 3 and Grade 6 portfolios newly accepted and evaluated in 2006, participation rates increased by 44% from 2005 to 2006. Among them, 65% were males and 35% were females. A total of 4,728 student portfolios were received.

The content area “Communication” was excluded from the scoring domain this year, leaving four content dimensions in Part I to assess progress on IEP skills and four performance dimensions in Part II to evaluate students’ generalization of skills. The technical quality of scores such as scorer reliability remains consistent from year to year. As in previous years, the inter-rater reliability is generally higher for Part I than Part II. Although inter-correlations between Part I and Part II scores were fairly high across all grades, statistical evidence is consistent with the two-dimensional construct of WAAS: The two parts are measuring different traits. The 2006 score distributions again reveal student growth from previous years. While the frequency of the highest score (4) for some of the Part II dimensions declined slightly from 2005 to 2006, overall scores increased. Progress on the content dimension improved significantly from 2000. To date, the 2006 WAAS produced the highest percentage of students meeting the performance standard since the program’s inception in 2000.

II. What is WAAS?

State assessment programs provide one method of determining student academic achievement. The Washington State Assessment System provides accountability for program and educational opportunities for all students. Alternate assessment is one component of Washington's assessment system.

The Washington Alternate Assessment System (WAAS) program was developed by the Washington Alternate Assessment Task Force and expanded by Advisory Panels in response to requirements of the Individuals with Disabilities Education Act of 1997: "The State has established goals for the performance of children with disabilities in the state that . . . are consistent, to the maximum extent appropriate, with other goals and standards for children established by the state." The alternate assessments are based on Washington's Essential Academic Learning Requirements (EALRs) in the content areas of Communication, Reading, Writing, Mathematics, and Science. The state has prepared extensions for the EALRs. This document describes the critical function of the EALRs, the access skills, instructional activities, and assessment strategies that are designed to assist special education staff members to link functional IEP skills to the EALRs, to provide access to the general education curriculum, and to measure student progress toward achieving the EALRs. The most current version of the EALR extensions document can be found at <http://www.k12.wa.us/SpecialEd>

A number of additional resources and documents can be downloaded from <http://www.k12.wa.us/SpecialEd/assessment.aspx>

The inclusion of students with disabilities in the assessment and accountability system is critical to ensure appropriate allocation of resources and learning opportunities for these students.

The Washington Alternate Assessment System was designed for a small percentage of the total school population. Students with disabilities are expected to take the Washington Assessment of Student Learning (WASL) tests, with or without necessary accommodations, unless the Individualized Education Program (IEP) team determines that the student is unable to participate on the WASL in one or more content areas. In this case, the IEP team may select the Washington Alternate Assessment System (WAAS) portfolio assessment.

Program Purpose

The Washington Alternate Assessment Task force—comprised of administrators, higher education personnel, teachers, and parents—determined the following two-fold purpose of the portfolio assessment:

- To provide an appropriate method of measuring progress on state goals and standards for students who are not able to access the WASL or any commercially available test, even with accommodations
- To ensure that students will be able to generalize the Individualized Education Program (IEP) skills to the maximum extent possible

The basic building block of the portfolio assessment is evidence of the student's performance and progress toward reaching IEP goals. Each of the entries in the portfolio documents two dimensions of learning: progress on IEP skills linked to the EALRs and student generalization of those skills.

Portfolio evidence should demonstrate participation in and progress toward IEP goals that are aligned to state standards (EALRs). In this way, evidence of progress on IEP skills is linked to the EALRs and can be linked to progress on state goals and standards.

Portfolio evidence should also show the extent to which a student can demonstrate and generalize the IEP skill linked to EALRs in the following ways:

- Use the IEP skill with appropriate modifications/adaptations, supports, or assistive technology in order to demonstrate all he or she knows and is able to do.
- Apply the IEP skill in a variety of settings and contexts in which the student is able to use learned skills. These places can include the classroom, other areas of the school, community settings, and home.
- Interact with non-disabled peers and others during IEP activities for the purpose of developing social relationships to enrich his or her life.
- Use self-determination skills in planning, monitoring, and evaluating IEP skill activities.

Target Population & Participation Rates

Federal guidelines indicate that states should develop alternate assessment participation guidelines so that approximately 1%–3% of the student population is eligible for an alternate assessment in each given year.

The number of portfolios submitted in 2006 increased mainly due to Grade 3 and Grade 6 participation, which were absent in 2005. In 2006, 4,728 students participated in the WAAS program. As shown in Table 2.1, the participation rate of the WAAS increased at each grade level except Grade 10. In 2002, a total of 427 portfolios were submitted. The number almost quadrupled in 2003 (1,646), rose to 2,589 in 2004, and rose to 3,279 in 2005. The number reached a new high in 2006.

Table 2.1 Number of Participants in the WAAS from 2002

Grade	Number of 2002 Participants	Number of 2003 Participants	Number of 2004 Participants	Number of 2005 Participants	Number of 2006 Participants
3					720
4		695	726	780	801
5			389	605	685
6					630
7		425	531	678	680
8		174	440	589	680
10		352	503	627	532
Total	427	1,646	2,589	3,279	4,728

III. PORTFOLIO DEVELOPMENT

The implementation of the WAAS portfolio is dependent on the interaction between the assessed student and the teacher or staff member who assists the student with portfolio construction. The teacher and the student must be cognizant of the components and types of evidence that are required and/or recommended for inclusion in the portfolio. The student must be able to demonstrate observable skills or to produce evidence to be included in the portfolio.

The teacher or staff member must be able to write measurable IEP goals or objectives that provide opportunities for the student to participate and progress in the general curriculum. Staff members must also be able to plan academic content-based activities and select one IEP skill linked to EALRs that will be measured in each content area entry. Additionally, the assessment team that collects data (on the student's progress on IEP skills over time and the ability of the student to generalize and use these skills) must possess a certain level of assessment literacy about how best to measure assessment targets and document student growth in IEP skills.

Two ongoing activities have been implemented to document and control the effects of teacher knowledge of WAAS portfolio procedural issues or assessment practices. Regional teacher training sessions are conducted in the fall of each school year. Workshops are conducted in several regions of the state in the fall each year. The three-hour workshops cover implementing WAAS procedures, writing measurable IEP goals, planning general education content-based activities, and collecting student performance data. The WAAS portfolio session materials are posted to the OSPI Web site, and members of the Special Education Assessment Leadership Team (SEALT) are trained to replicate the WAAS workshops for teachers who missed an OSPI session. Participant surveys are conducted at the regional workshops to gauge perceptions of changes in instruction and assessment practices and to determine other training needs.

Research, data analyses, and reviews of the WAAS portfolios for a sample of copied portfolios each year has been the second strategy to inform us about the technical adequacy of the WAAS portfolio and to guide professional development for future WAAS workshops (Johnson & Arnold, 2004). Johnson addresses the following research questions in a 2004 report:

1. Does the WAAS include tasks from the extended benchmarks for the EALRs? If so, what kind and how frequently are they used?

2. How many low scores might be attributable to procedural issues with the portfolio?
3. In what way might we establish external validity of the WAAS in future administrations?

The findings of this study, perception surveys, and other studies were used to determine the research agenda for the 2006 WAAS portfolios review and to revise WAAS portfolio workshops to address professional learning needs.

Construct Validity

Common approaches to investigate construct validity of an assessment are to examine inter-dimension correlations and factor structure. The correlation method provides information about the relationships among the test dimensions. Similarly, the factorial method explores the structure of an assessment through the correlations among the sub-scores.

Tables 3.1 to 3.7 show the correlations among the sub-scores for all the grades. The correlations within Part I and within Part II are generally higher than the correlations between Part I and Part II. This result supports the two-part construct of the WAAS. The correlation matrices suggest that the Part II dimensions are different measures from the content progress. It is therefore reasonable to evaluate students' portfolios separately on their content progress and on their performance dimensions.

Table 3.1 Grade 3 Sub-score Correlations

	Part I		Part II			
	Reading	Math	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Reading	1.000					
Math	.335	1.000				
Modifications/ Adaptations	.239	.205	1.000			
Settings/ Contexts	.290	.215	.578	1.000		
Social Relations	.311	.212	.428	.564	1.000	
Self- Determination	.309	.139	.310	.400	.430	1.000

Table 3.2 Grade 4 Sub-score Correlations

	Part I			Part II			
	Reading	Writing	Math	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Reading	1.000						
Writing	.484	1.000					
Math	.482	.575	1.000				
Modifications/ Adaptations	.328	.312	.297	1.000			
Settings/ Contexts	.348	.310	.289	.648	1.000		
Social Relations	.360	.326	.309	.534	.589	1.000	
Self- Determination	.251	.232	.223	.384	.446	.436	1.000

Table 3.3 Grade 5 Sub-score Correlations

	Part I			Part II			
	Reading	Math	Science	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Reading	1.000						
Math	.602	1.000					
Science	.320	.452	1.000				
Modifications/ Adaptations	.376	.342	.323	1.000			
Settings/ Contexts	.388	.342	.312	.591	1.000		
Social Relations	.344	.305	.380	.483	.602	1.000	
Self- Determination	.276	.235	.290	.371	.412	.466	1.000

Table 3.4 Grade 6 Sub-score Correlations

	Part I		Part II			
	Reading	Math	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Reading	1.000					
Math	.528	1.000				
Modifications/ Adaptations	.354	.379	1.000			
Settings/ Contexts	.419	.384	.602	1.000		
Social Relations	.426	.385	.501	.637	1.000	
Self- Determination	.363	.347	.410	.459	.509	1.000

Table 3.5 Grade 7 Sub-score Correlations

	Part I			Part II			
	Reading	Writing	Math	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Reading	1.000						
Writing	.570	1.000					
Math	.567	.543	1.000				
Modifications/ Adaptations	.393	.373	.344	1.000			
Settings/ Contexts	.403	.394	.356	.609	1.000		
Social Relations	.408	.411	.377	.484	.602	1.000	
Self- Determination	.327	.340	.299	.388	.467	.531	1.000

Table 3.6 Grade 8 Sub-score Correlations

	Part I			Part II			
	Reading	Math	Science	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Reading	1.000						
Math	.648	1.000					
Science	.348	.276	1.000				
Modifications/ Adaptations	.319	.338	.255	1.000			
Settings/ Contexts	.325	.353	.306	.617	1.000		
Social Relations	.334	.384	.258	.521	.588	1.000	
Self- Determination	.222	.256	.267	.381	.403	.487	1.000

Table 3.7 Grade 10 Sub-score Correlations

	Part I				Part II			
	Reading	Writing	Math	Science	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Reading	1.000							
Writing	.663	1.000						
Math	.685	.624	1.000					
Science	.547	.645	.646	1.000				
Modifications/ Adaptations	.460	.501	.457	.516	1.000			
Settings/ Contexts	.450	.544	.451	.527	.730	1.000		
Social Relations	.430	.518	.431	.475	.613	.688	1.000	
Self- Determination	.299	.453	.331	.434	.504	.585	.486	1.000

An exploratory factor analysis was performed to further investigate the nature of the inter-correlations across the score dimensions. A principle components analysis with Orthogonal Varimax Rotation was conducted using SAS (Statistical Analysis Software, v 9.1). The number of factors was determined using three criteria: eigenvalues greater than 1, a scree test for the eigenvalues, and finding the solution in which approximately 40% of the variance was explained. The result was a one-factor solution for all grades, where 45%–59% percent of the variance was explained. Table 3.8 shows the factor loadings from the component matrix for the factor analysis. It also lists the total variance explained by the extracted factors. This result is consistent with the analysis of the inter-correlations.

While both analyses suggest that 2006 WAAS scores may be correlated, clearly Part I and Part II are separate dimensions of performance on the WAAS as a whole. Based on the a priori hypothesis that there should be two distinct factors, a two-factor solution was also investigated for all grades. As shown in the “Factor 2” columns of Table 3.8, the factor loadings of the two-factor solution show that the content progress scores (Part I) and performance scores (Part II) exhibit a distinguished loading pattern on the second factor.

Table 3.8 Factor Loadings

	Score Dimension	Factor 1	Factor 2
Grade 3			
Part I	Reading	.28	.73
	Math	.06	.86
Part II	Modifications/Adaptations	.75	.10
	Settings/Contexts	.83	.14
	Social Relations	.77	.18
	Self-Determination	.65	.17
Variance Explained		45.20%	16.81%

	Score Dimension	Factor 1	Factor 2
Grade 4			
Part I	Reading	.26	.73
	Writing	.18	.83
	Math	.15	.84
Part II	Modifications/Adaptations	.79	.22
	Settings/Contexts	.84	.20
	Social Relations	.77	.25
	Self-Determination	.69	.11
Variance Explained		48.45%	17.19%
Grade 5			
Part I	Reading	.23	.80
	Math	.13	.89
	Science	.31	.61
Part II	Modifications/Adaptations	.71	.29
	Settings/Contexts	.79	.25
	Social Relations	.80	.22
	Self-Determination	.71	.11
Variance Explained		48.29%	15.01%
Grade 6			
Part I	Reading	.27	.83
	Math	.23	.85
Part II	Modifications/Adaptations	.77	.19
	Settings/Contexts	.83	.22
	Social Relations	.79	.26
	Self-Determination	.68	.25
Variance Explained		54.13%	14.15%

	Score Dimension	Factor 1	Factor 2
Grade 7			
Part I	Reading	.27	.80
	Writing	.27	.79
	Math	.20	.83
Part II	Modifications/Adaptations	.73	.26
	Settings/Contexts	.82	.24
	Social Relations	.78	.28
	Self-Determination	.73	.17
Variance Explained		52.51%	14.67%
Grade 8			
Part I	Reading	.14	.89
	Math	.20	.84
	Science	.27	.51
Part II	Modifications/Adaptations	.76	.23
	Settings/Contexts	.79	.25
	Social Relations	.78	.25
	Self-Determination	.71	.11
Variance Explained		47.21%	15.70%
Grade 10			
Part I	Reading	.20	.85
	Writing	.40	.76
	Math	.21	.86
	Science	.40	.71
Part II	Modifications/Adaptations	.77	.33
	Settings/Contexts	.84	.31
	Social Relations	.76	.31
	Self-Determination	.77	.16
Variance Explained		58.63%	13.28%

Content Validity

Content-related validity evidence addresses the extent to which the assessment tasks adequately align to the standards intended as the focus of assessment. The work of the content workgroups and development of EALR extensions (described in further detail in the next section) provides evidence that the WAAS measures the intended content standard or access skills objectives. Content experts (listed in Appendix C) reviewed every mastery objective to ensure alignment to, and appropriate representation of, the underlying objective identified by the test examiner. These experts provided feedback to test examiners regarding how the mastery objective can be improved and whether alignment is an issue.

IV. CONTENT WORKGROUPS & DEVELOPMENT OF EALR EXTENSIONS

Content Workgroups

The content workgroups were formed to build on the Essential Academic Learning Requirements (EALR) Extension Curriculum Guide created by a group of state stakeholders in the spring of 2000. In 2000, the curriculum group developed a document with suggested learning activities and functional access skills linked to content area standards. As part of a team of Washington educators, the participants began creating a curriculum guide linked to the EALRs, and specifically linked to the Grade Level Expectations (GLEs) for Reading, Writing, and Mathematics. This document is different in design and purpose from the EALR Extension Curriculum Guide. The Curriculum Alignment Document is intended to illustrate examples of how students with the most significant disabilities can access the general curriculum for learning.

The workgroup convened for three days, February 8–10, 2006. Day one consisted of training to begin the change process and discussion about what learning may look like for students who access the general curriculum. The following topics were covered:

- Who the Kids are Who Take Alternate Assessments
- History of Special Education Curriculum (Developmental, Functional, Social Inclusion, and Access to the General Curriculum)
- Universal Design for Learning
- Planning for Instruction (Four-step process)

- Curriculum Frameworks Form

The broad state standard (EALR) must first be identified before the corresponding grade-level expectation (GLE) can be identified. These steps are distinct from first identifying the IEP goal or objective and then linking back to the standard. In this model, the grade-level standards are the foundation for instructional decision making.

The next step is to identify or to design a general education instructional activity that moves students toward acquisition of the standard. The general education activity is broken down into instructional steps. A description of the type of student and potential IEP skills that can be addressed within the instruction are defined.

The workgroups then identify how a student with the most significant disabilities can participate in each of the instructional steps that correspond to the “progress” dimension of the rubric. The student’s participation in each of the instructional steps is then specified, including all supports (e.g., adaptations, modifications, assistive technology) the student will need. Possible evidence for the assessment can be designated, and all rubric dimensions should be considered. Opportunities to embed other IEP goals or objectives can then be identified to emphasize “functional” skills as a focus of instruction.

While this process is a preliminary step to initiate discussions about grade-level curriculum in an extensive development process, the work completed by this group could lead to a resource document that supplements the curriculum resources currently available on the OSPI Web site. The goal is to develop an example for each EALR at the grade level assessed and, using the attached format, show how a student with significant disabilities can access the curriculum and produce evidence for assessment.

Workgroup days two and three consisted of break-out sessions by content area and by grade-level clusters for all grades. Each grade-level cluster first identified the characteristics of a child with significant disabilities and wrote a student vignette describing the student’s strengths, weaknesses, or learning styles. The sample curriculum alignment document was then designed around the student description, using the four-step process for planning and the curriculum alignment form.

Membership

It is important to include participants who have aligned classroom instruction to the GLEs and educators who are familiar with students with significant disabilities to create this resource.

The workgroups were initially designed to include one-third special educators, one-third general educators, and one-third content experts for each of Reading, Writing, and Math.

The actual makeup of the workgroups consisted of 34 special educators and 10 general educators/content experts.

Application and Implementation

The revised curriculum alignment documents were presented at all of the Summer Institutes. Comments and feedback were collected about the presentation and the curriculum alignment form.

ILSSA continues to revise the current draft examples. The content workgroups are ongoing and, when convened again, will resume work on the revised examples and develop additional examples for incomplete EALRs. The goal is to provide these resources to all teachers through regional training workshops and to disseminate these materials through the OSPI Web site.

V. SUMMER SCORING INSTITUTE

There are five scoring dimensions divided into two parts. One dimension is scored on specific content area sections of the portfolio, and four dimensions are scored across the entire portfolio. Part I scores for Progress on IEP skills are determined by evidence in separate portfolio entries for Reading, Writing, Mathematics, and Science. Part II scores for Student Generalization of Skills in four dimensions are determined by examining evidence across the entire portfolio. The content area Part I score is added to the total of the four dimension scores in Part II to obtain a total score for the content area. This produces two total scores for grades 3 and 6 (Reading and Mathematics), three separate total scores (Reading, Writing, and Mathematics) for the students in grades 4 and 7, three total scores (Reading, Mathematics, and Science) for grades 5 and 8, and four scores for grade 10 (Reading, Writing, Mathematics, and Science). Table 5.1 lists the various scoring dimensions and the applicable grades. Appendix F includes the scoring sheet used during the process. The scoring sheet also provides definitions for each score category.

Table 5.1 Contents and Dimensions of the 2006 WAAS Scoring

	IEP Skills	Grades Assessed
Part I	Mathematics	All
	Reading	All
	Science	5, 8, 10
	Writing	4, 7, 10
Part II	Modifications and Adaptations	All
	Settings and Contexts	All
	Social Relations	All
	Self-Determination	All

The portfolio rangefinding and scoring occurred over a one-week period in June (June 12–16). During the week, a group of representatives from Pearson Educational Measurement (PEM) were led by OSPI staff to review the anchor portfolios and prepare for scoring. The PEM staffs have many years of experience scoring the alternate assessment. The anchor portfolios, taken from previous years, exemplified score points for the rubric. Potential table leaders were trained to use the rubric using the exemplar portfolios. The potential table leaders are reliable and experienced scorers from previous years. Potential table leaders were chosen after successfully scoring a qualifying portfolio. Table leaders had to demonstrate an exact plus an adjacent match rate of 90% or better in order to qualify as a table leader. That week, fifteen portfolios were scored during rangefinding. PEM and OSPI personnel reviewed all the scored portfolios, and four portfolios were selected for training, two portfolios were selected for qualifying portfolios, and nine portfolios were chosen to be used as daily validation papers. Scoring summaries and annotations were written to accompany the training sets for the Portfolio Scoring Institute.

Interested teachers completed an application to become WAAS portfolio scorers. OSPI used three criteria to select scorers: attendance at a workshop to implement the portfolio, participation in constructing a portfolio, and explanation of why the scorer wanted to score portfolios.

Scoring occurred the next week. OSPI assigned scorers to each table, and table leaders were randomly assigned to the tables. PEM leadership staffs served as scoring trainers. Some teachers participated during both weeks of rangefinding and scoring, but most scorers during the second week were new and the process was repeated for the second week. The first day of each week was dedicated as a full day of training. During the second week, returning teachers did not participate in the training but continued with scoring.

Potential scorers were trained to use the rubric using the exemplar portfolios. Teachers were trained to score independently using three exemplar portfolios. OSPI and PEM facilitated discussion upon completion of scoring. When OSPI and PEM concluded that all teachers were properly trained, scoring procedures were reviewed. After training, the teachers scored a qualifying portfolio. Teachers who did not qualify after the first qualifying session were given an opportunity to score a second qualifying portfolio. All scorers were trained and given a list of those things that should not bias their scoring.

Appendix E describes the scoring procedures. All scorers were considered to be first scorers until sufficient rater reliability information was obtained. The first scorer logged the portfolio and listed the table colors on the Sign Out/Sign In sheet. Scorers were not allowed to select a portfolio from their own school district or from school districts of their tablemates to score. Scorers used the white Scoring Summary Sheets to record scores. Once scored, the sheet was placed upside down in the portfolio pocket. Scorers wrote their badge number and table color on the back of the sheet. Scorers then completed the portfolio checklist and placed it in the inside cover of the binder. After scoring the portfolio, scorers placed a check mark on the Sign Out/Sign In sheet for that portfolio. The portfolio was then placed at Station Number Two.

All table leaders were second scorers on days 1 and 2. In order to be qualified as a second scorer on Day 3, scorers were required to exhibit the following reliability evidence: pre-scored qualifying set at 75% exact plus adjacent scores, one validation paper at least 80% exact or adjacent match, and inter-rater reliability statistics by the end of Day 2 of at least 90% exact or adjacent

match. Table leaders were consulted about the level of independence and accuracy of the first scorer before he or she was assigned to be a second scorer.

Qualified second scorers picked up a portfolio at Station Number Two. Scorers logged the portfolio and listed their table color on the Sign Out/Sign In sheet. Second scorers utilized a yellow Scoring Summary sheet and placed the sheet in the front portfolio pocket, noting his or her badge number on the back of the sheet. After scoring the portfolio, scorers put a check mark on the Sign Out/Sign In sheet for that portfolio. The portfolio was placed at Station Number Three.

Portfolios were blind double-scored on Day 2. The leadership team and table leaders were second scorers. Table leaders also conducted at least one blind read or a “backread” of the scorers at their tables. Table leaders and scorers completed a validation portfolio each afternoon. The leadership team also reviewed all reliability statistics, including validation agreement, each afternoon.

The first and second scoring summary sheets were compared at Station Number Three. For grades 3–8, if any scores were adjacent scores, the scores of the first scorer were used as the student’s final score. In Grade 10, the adjacent scores were added as the student’s final score. If scores on a dimension were not adjacent, a third scoring was completed by a member of the leadership team. The same process was followed for Grade 10 if the total scores for a subject resulted in different performance levels. For scores that were not adjacent, the sections requiring a third read were noted, and the portfolio was placed on the third read table. Discrepant dimensions only, rather than the entire portfolio, were reviewed and rescored. The final scores were copied onto the NCR Scoring Summary sheet (three-part form) to be returned to school districts. Final scores were entered on the student’s WAAS demographic form for reporting by PEM.

For portfolios with “backread” scores, the backread scores were treated as the second and final read. Backread scores were written on NCR paper with a note on the white sheet designating it as a final score. Then “backread” was noted on the back of the first reader’s scoring summary sheet next to the scorer’s number, and the backread scoring summary was placed in the backread box.

Inter-rater reliability statistics were calculated each day. Whole-group recalibration and training occurred for any scoring dimension statistic that was discrepant with other dimensions or that deviated greatly from previous years. Individual scorer retraining occurred for those scorers with less than 80% exact or adjacent scores on validation portfolios.

Table leaders and scorers were asked each day to score a validation portfolio, which was a portfolio that had been previously scored by OSPI/PEM portfolio leadership staff. The leadership staff scored the portfolios and looked for scorers or table leaders who had less than 80% exact matches and adjacent scores. Scorers with lower scores were retrained.

To ensure that students received accurate scoring judgments, these two procedures were followed:

- Second reads were performed on all Day 2 portfolios.
- Table leaders “backread” at least one portfolio from each scorer at their table daily.

Reliability of scorers was monitored by comparing scorer performance relative to those conducting second reads and by reviewing their performance on validation portfolios.

VI. TECHNICAL QUALITIES OF SCORING

In this section, qualities of the 2006 WAAS portfolio scores are examined through the analysis of variance between and within scorers. As described in the scoring section, the between-scorer difference was closely monitored during the scoring process. An inter-rater agreement index was regularly computed to control between-scorer variation. Also examined in this section is the Alpha Coefficient for each grade, which provides information about the internal consistency of the score dimensions.

Inter-Rater Agreement

Inter-rater agreement is an important source of evidence for the reliability of test scores. When two trained judges agree with the score given to a student’s work, it supports the concept that this is the “correct” score for that student’s work. The percent of agreement between scorers is examined to determine the degree to which judges gave equivalent scores to the same student work. The reliability of scoring is described in the following ways:

1. Differences between the scores from the first and second reader for each scorer each day was monitored. Tables 6.1–6.8 summarize the extent of agreement between the first and second scorers for each scoring dimension in each grade over five days of scoring. As shown in Table 6.8, the overall percentage of exact agreement or of adjacent scores for all the four content areas (Part I) is quite high (ranging from 87% to 91%). This is also true when the analysis is performed separately for each grade (tables 6.1–6.7). On the other hand, Part II scores generally exhibit lower inter-rater agreement relative to Part I scores. This is especially true for the modifications/adaptations and self-determination scores, where perfect agreement rarely exceeds 70% and the combination of perfect agreement and adjacent scores is always below 95%. The difference in score agreement between Part I and Part II may be explained by different interpretations of the Part II scoring definition by raters and limited evidence for Part II scores. These results are, however, consistent with those from 2002 and 2005. While variation in agreement across scoring dimensions does exist, these percentages of agreement appear to be fairly reasonable.

Table 6.1 Percentage Agreement between First and Second Scorers for Grade 3

Extent of Agreement	Part I			Part II			
	Reading	Writing	Math	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Perfect Agreement	71.6	NA	71.6	55.6	44.4	61.7	67.9
Adjacent	12.3	NA	12.3	33.3	43.2	30.9	22.2
Non-Adjacent	16.0	NA	16.0	11.2	12.4	7.4	9.8

Table 6.2 Percentage Agreement between First and Second Scorers for Grade 4

Extent of Agreement	Part I			Part II			
	Reading	Writing	Math	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Perfect Agreement	73.9	75.0	78.4	39.8	48.9	47.7	63.6
Adjacent	13.6	18.2	12.5	39.8	40.9	46.6	28.4
Non-Adjacent	12.6	6.8	9.0	20.4	10.2	5.6	8.0

Table 6.3 Percentage Agreement between First and Second Scorers for Grade 5

Extent of Agreement	Part I			Part II			
	Reading	Math	Science	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Perfect Agreement	75.7	83.8	75.7	48.6	48.6	55.4	63.5
Adjacent	13.5	9.5	10.8	36.5	41.9	40.5	23.0
Non-Adjacent	10.8	6.8	13.6	14.8	9.4	4.0	13.6

Table 6.4 Percentage Agreement between First and Second Scorers for Grade 6

Extent of Agreement	Part I			Part II			
	Reading	Writing	Math	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Perfect Agreement	82.6	NA	72.5	49.3	56.5	53.6	79.7
Adjacent	7.2	NA	17.4	34.8	30.4	44.9	10.1
Non-Adjacent	10.2	NA	10.2	16.0	13.0	1.4	10.2

Table 6.5 Percentage Agreement between First and Second Scorers for Grade 7

Extent of Agreement	Part I			Part II			
	Reading	Writing	Math	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Perfect Agreement	62.2	71.6	71.6	44.6	54.1	55.4	67.6
Adjacent	17.6	12.2	14.9	40.5	39.2	35.1	13.5
Non-Adjacent	20.2	16.2	13.6	14.8	6.8	9.4	19.0

Table 6.6 Percentage Agreement between First and Second Scorers for Grade 8

Extent of Agreement	Part I			Part II			
	Reading	Math	Science	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Perfect Agreement	81.6	75.0	78.9	48.7	42.1	43.4	63.2
Adjacent	9.2	13.2	11.8	36.8	46.1	48.7	25.0
Non-Adjacent	9.2	11.8	9.2	14.4	11.8	7.8	11.8

Table 6.7 Percentage Agreement between First and Second Scorers for Grade 10

Extent of Agreement	Part I				Part II			
	Reading	Writing	Math	Science	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Perfect Agreement	79.5	75.0	78.3	74.5	44.1	48.8	50.7	58.0
Adjacent	12.1	13.8	12.5	13.1	39.3	41.1	38.2	25.7
Non-Adjacent	8.4	11.1	9.2	12.5	16.5	10.2	11.1	16.3

Table 6.8 Overall Percentage Agreement between First and Second Scorers

Extent of Agreement	Part I				Part II			
	Reading	Writing	Math	Science	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Perfect Agreement	77.1	74.4	76.8	74.9	45.9	48.8	51.5	62.1
Adjacent	12.3	14.4	13.0	12.8	38.3	40.7	39.7	23.3
Non-Adjacent	10.6	11.2	10.2	12.4	15.8	10.4	8.8	14.6

2. The results from daily scoring of validation portfolios was monitored. Each day, scorers scored a previously scored “validity” portfolio. The results from these scores were captured and reviewed daily. During the two weeks, scores were captured for 225 reads (123 for Grade 5, and 51 for Grade 7 and Grade 10 each) of validity portfolios. As in the regular portfolio scoring, there are five scores (one content and four performance dimensions). Table 6.9 shows the frequency distribution of the number of scores that reached exact or adjacent agreement between scores and experts. Approximately 94.7% of the 225 readings had at least four out of five scores that were exact or adjacent. A detailed analysis of the agreement between scorers and experts on validity portfolios is presented in tables 6.10 and 6.12. The rate of agreement was examined separately for each grade and content/dimension.

Table 6.9 Frequency Distribution of the Number of Scores Reached at Least Adjacent Agreement on Validity Portfolios

Number of Scores (out of 5) That Reached Exact or Adjacent Agreement on the Validity Portfolios	Frequency	Percentage
All 5 Scores Reached Exact/Adjacent Agreement	148	65.8%
4 out of 5 Scores Reached Exact/Adjacent Agreement	65	28.9%
3 out of 5 Scores Reached Exact/Adjacent Agreement	8	3.6%
2 out of 5 Scores Reached Exact/Adjacent Agreement	4	1.8%
1 out of 5 Scores Reached Exact/Adjacent Agreement	0	0%
0 out of 5 Scores Reached Exact/Adjacent Agreement	0	0%
Total	225	100.0%

Table 6.10 Extent of Agreement between Scorers and Experts on Validity Portfolios: Grade 5

Extent of Agreement	Part I			Part II			
	Reading	Math	Science	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Perfect Agreement	91.1	82.1	95.9	37.4	65.0	39.8	65.9
Adjacent	8.9	9.8	2.4	62.6	35.0	57.8	32.5
Non-Adjacent	0.0	8.1	1.6	0.0	0.0	2.4	1.6

Table 6.11 Extent of Agreement between Scorers and Experts on Validity Portfolios: Grade 7

Extent of Agreement	Part I			Part II			
	Reading	Writing	Math	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Perfect Agreement	72.5	72.5	64.7	94.1	92.2	90.2	92.2
Adjacent	23.5	27.5	33.3	5.9	7.8	9.8	7.8
Non-Adjacent	4.0	0.0	2.0	0.0	0.0	0.0	0.0

Table 6.12 Extent of Agreement between Scorers and Experts on Validity Portfolios: Grade 10

Extent of Agreement	Part I				Part II			
	Reading	Writing	Math	Science	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Perfect Agreement	98.0	100.0	98.0	100.0	76.5	41.2	68.6	94.1
Adjacent	2.0	0.0	2.0	0.0	17.6	45.1	27.4	5.9
Non-Adjacent	0.0	0.0	0.0	0.0	5.9	13.7	4.0	0.0

Coefficient Alpha

Coefficient Alpha is a score reliability index of internal scale consistency or homogeneity. Alpha can be estimated from scores obtained on one occasion and is appropriate when a score is intended to measure a single trait. Coefficient Alpha was computed using the formula below:

$$\alpha = \left[\frac{n}{n-1} \right] \left[1 - \frac{\sum_{i=1}^n \sigma^2(x_i)}{\sigma^2(X)} \right]$$

where n is the number of items or components,

$\sigma^2(x_i)$ is the item or component score variance ($i = 1, 2, \dots, n$), and

$\sigma^2(X)$ is the total score variance

The value of Alpha is affected by the number of components comprising a score, the variance of the individual components, and the total score variance. In the context of the WAAS scores, relatively higher values of Alpha will result when the total scores have greater variability and/or the score variance across the individual components is very small (i.e., internally consistent). Table 6.13 shows the Coefficient Alpha of the total score for each subject at each grade. The Coefficient Alpha based on Part II scores alone is also presented. The resulting values of Coefficient Alphas were fairly high for the relatively small number of components (four or five in this analysis), which may be attributed to the small variance within each component.

Table 6.13 Coefficient Alpha for Total Scores and Part II Scores

Grade	Total Score				Part II Score
	Reading + Part II	Writing + Part II	Math + Part II	Science + Part II	
3	.75	—	.70	—	.75
4	.78	.77	.77	—	.80
5	.78	—	.76	.77	.78
6	.81	—	.80	—	.80
7	.80	.81	.79	—	.80
8	.77	—	.78	.76	.79
10	.83	.86	.84	.85	.85

VII. SUMMARY OF STUDENT PERFORMANCE

Of the 4,728 students who submitted portfolios, 1,650 (35%) reported as female and 3,062 (65%) reported as male. Sixteen (0.3%) students did not report their gender.

The results reported in Table 7.1 are based on the data captured in Box 20 of the demographic sheet and are the results reported to parents, schools, districts, and state. This represents the final post-record change data for the state. Table 7.1 provides a summary of the percentage of students earning each of the scores in each dimension scored. Like previous years, Part I scores are generally higher than Part II scores. However, the difference in score distributions between Part I and Part II seems to have increased this year.

Raising Content Progress (Part I) Scores

The content scores were generally lower in previous years. Many students were given a score of 1 or 2 for most content domains. In 2006, however, there was a significant increase in the number of students receiving scores of 3 or 4. In 2001, only 5.2% of the portfolios in Mathematics were scored at level 4. Percentages were low in the other domains as well. The distributions of the content scores gradually changed from year to year, with more students earning higher scores of 3 or 4. In 2005, the percentages of the highest score for all grades and contents ranged from 33.2% to 46.0%. This year, the percentages of the highest score range from 44.2% to 58.8%.

Improvement of Part II Scores

Scores on the four performance dimensions (Part II) also show general improvement from the previous year. In 2005, fewer students were awarded a score of 3 or 4 on the performance dimensions, especially on Modifications. The percentage of highest score for Social Relations was consistently below 7% across all grade levels. In 2006, at least 8.3% (Grade 10) of the students scored 4 on the same dimension. Grade 10 has a wider score range this year because of the summation rule applied to this specific grade. If we examine the score distributions for grades 3–6 (of which the scale remains consistent with the previous years), the percentage of students receiving the highest score on Part II is higher than 2005 across all dimensions.

Table 7.1 2006 WAAS Score Distributions

Score Point	Part I				Part II			
	Reading	Writing	Math	Science	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Grade 3								
0	2.1		8.6		1.3	1.3	1.3	1.3
1	10.0		11.8		18.1	18.3	24.3	44.3
2	15.6		13.6		27.2	30.1	43.5	13.8
3	13.6		13.1		25.7	31.3	17.5	7.4
4	58.8		52.9		27.8	19.0	13.5	33.3
Grade 4								
0	3.6	5.9	7.9		1.4	1.4	1.4	1.4
1	10.9	14.2	13.5		15.6	14.1	26.3	40.2
2	16.9	15.6	14.5		23.7	30.0	37.0	13.7
3	10.5	14.6	14.0		29.0	32.2	20.7	10.2
4	58.2	49.7	50.2		30.3	22.4	14.6	34.5
Grade 5								
0	9.9		13.6	11.8	2.2	2.2	2.2	2.2
1	12.4		14.3	18.0	26.1	20.3	26.3	48.5
2	13.1		12.1	8.2	24.1	28.0	41.8	12.7
3	9.6		11.7	13.9	21.5	31.7	19.3	11.4
4	54.9		48.3	48.2	26.1	17.8	10.5	25.3

Values in cells represent percentage.

Table 7.1 2006 WAAS Score Distributions (Continued)

Score Point	Part I				Part II			
	Reading	Writing	Math	Science	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Grade 6								
0	3.5		5.4		2.1	2.1	2.1	2.1
1	13.8		12.4		25.4	21.9	31.8	38.7
2	15.7		17.3		20.2	24.6	30.6	14.1
3	12.4		12.7		24.0	27.8	19.5	6.7
4	54.6		52.2		28.4	23.7	16.0	38.4
Grade 7								
0	6.2	4.9	6.3		2.4	2.4	2.4	2.4
1	14.4	14.6	14.9		22.4	19.9	31.2	40.9
2	11.5	11.3	10.2		25.6	26.0	35.4	14.3
3	12.8	20.4	12.7		23.2	29.6	19.7	11.3
4	55.2	48.8	56.0		26.5	22.2	11.3	31.2
Grade 8								
0	13.4		13.2	9.7	1.6	1.6	1.6	1.6
1	8.5		11.8	13.1	24.1	15.6	26.3	36.0
2	11.8		8.5	9.3	21.9	25.7	32.4	13.5
3	15.0		13.7	15.6	21.8	31.6	21.0	9.6
4	51.3		52.8	52.4	30.6	25.4	18.7	39.3

Values in cells represent percentage.

Table 7.1 2006 WAAS Score Distributions (Continued)

Score Point	Part I				Part II			
	Reading	Writing	Math	Science	Modifications/ Adaptations	Settings/ Contexts	Social Relations	Self- Determination
Grade 10								
0	11.5	8.7	12.4	10.9	5.5	5.5	5.5	5.5
2	8.1	12.0	10.2	13.7	12.6	7.5	15.0	32.9
3	3.2	5.1	3.0	2.8	11.8	7.3	12.2	12.8
4	8.3	6.4	7.1	5.8	10.0	11.7	24.6	7.5
5	1.9	3.0	1.5	1.5	11.3	12.2	16.4	5.8
6	7.7	14.3	9.0	12.2	13.4	14.3	8.1	6.2
7	7.1	6.0	8.3	8.8	16.4	21.4	10.0	7.0
8	52.3	44.6	48.5	44.2	19.2	20.1	8.3	22.4

Values in cells represent percentage.

VIII. PERFORMANCE ON THE WAAS RELATIVE TO STANDARDS

Federal legislation and regulations for ESEA and IDEA reauthorization require states to report results for all students assessed using general assessments and alternate assessments relative to the same grade level academic content and achievement standards. The office of Superintendent of Public Instruction established four levels of performance based on alternate achievement standards on the WAAS assessments in the fall of 2003.

The standards for the portfolio were set using performance-level descriptors (see Appendix G) and copies of portfolios. The score needed for each performance level is based on the total score of the portfolio and a minimum score for Part I. To determine if a student meets standards, the total score is first determined by adding the score for each Part I score (progress on IEP skills for the content area) to the total score for the Part II dimensions (Modifications/Adaptations, Settings/Context, Social Relationships, and Self-Determination). The second part of the decision rule specifies a required minimum score on Part I (progress on IEP skills). Table 8.1 summarizes the decision rule for grades 3–8, whereas Table 8.2 summarizes the decision rule for Grade 10. Performance at level 3 or 4 is considered to be meeting standard. Portfolios with insufficient evidence were reported separately as IE and were not reported in one of the performance levels.

Table 8.1 Decision Rule for Determining Level of Performance on WAAS Portfolio (Grades 3–8)

	Points on Part II <i>Numbers in body of table are total points for (Part I) + (Part II)</i>																		
	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		
Part I Score = 1	5	Level 1					8	Level 2									17		
Part I Score = 2	6	Level 1				8	Level 2					11	12	Level 3					18
Part I Score = 3	7	7	8	Level 2			11	12	Level 3				15	16	Level 4			19	
Part I Score = 4	8	Level 2			11	12	Level 3				15	16	Level 4				20		

Table 8.2 Decision Rule for Determining Level of Performance on WAAS Portfolio (Grade 10)

	Points on Part II <i>Numbers in body of table are total points for (Part I) + (Part II)</i>																										
	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32		
Part I Score = 2	10	Level 1				14	15	Level 2																			34
Part I Score = 3	11	Level 1			14	15	Level 2						22	23	Level 3										35		
Part I Score = 4	12	Level 1		14	15	Level 2					22	23	Level 3							30	31	Level 4				36	
Part I Score = 5	13	14	15	Level 2			22	23	Level 3								30	31	Level 4					37			
Part I Score = 6	14	15	Level 2		22	23	Level 3						30	31	Level 4						38						
Part I Score = 7	15	Level 2			22	23	Level 3							30	31	Level 4					39						
Part I Score = 8	16	Level 2				22	23	Level 3								30	31	Level 4						40			

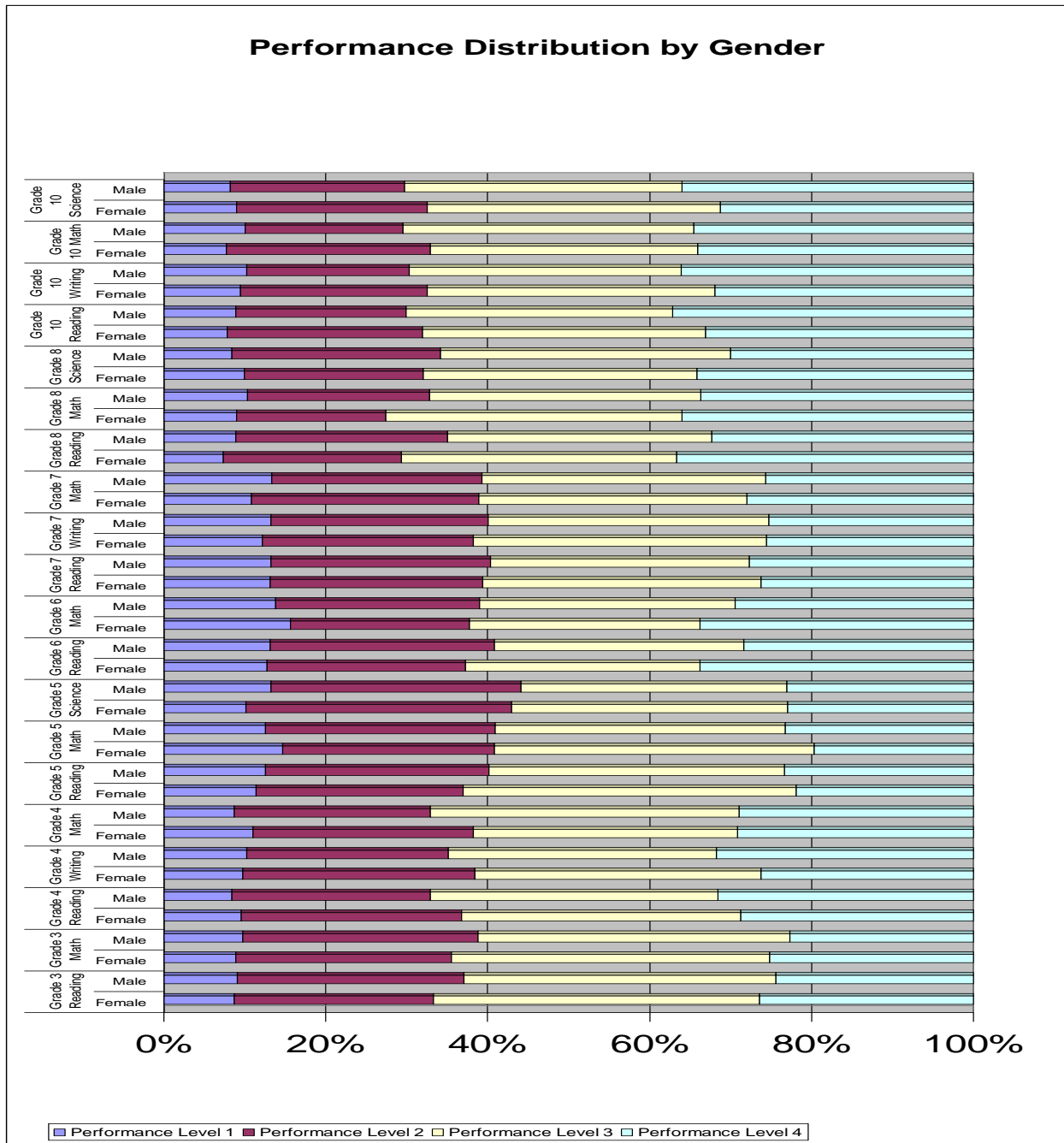
Table 8.3 shows the percentage of students achieving standards on the 2006 WAAS for each grade and content area. The achievement standards reported here are for the WAAS assessments and should not be compared to the results or standards for students taking the WASL.

Table 8.3 Percentage of Students by Performance Level on 2006 WAAS

		Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 10
Reading								
N count		720	801	685	630	680	680	532
Level 1	Not Meeting Standard	8.9%	8.8%	12.2%	13.0%	13.2%	8.5%	8.6%
Level 2		27.1%	25.4%	26.9%	26.6%	26.8%	24.5%	22.0%
Level 3	Meeting Standard	39.0%	35.1%	38.1%	30.1%	32.8%	33.3%	33.8%
Level 4		25.0%	30.7%	22.9%	30.3%	27.3%	33.8%	35.7%
Writing								
N count			801			680		532
Level 1	Not Meeting Standard		10.1%			12.9%		9.9%
Level 2		26.1%			26.5%		21.1%	
Level 3	Meeting Standard		33.8%			35.1%		34.4%
Level 4		30.0%			25.5%		34.6%	
Math								
N count		720	801	685	630	680	680	532
Level 1	Not Meeting Standard	9.6%	9.5%	13.3%	14.4%	12.4%	9.8%	9.2%
Level 2		28.3%	25.2%	27.5%	24.2%	26.7%	21.2%	21.5%
Level 3	Meeting Standard	38.8%	36.2%	37.2%	30.4%	34.2%	34.6%	35.0%
Level 4		23.4%	29.1%	22.0%	31.0%	26.7%	34.4%	34.3%
Science								
N count				685			680	532
Level 1	Not Meeting Standard			12.1%			9.1%	8.5%
Level 2		31.6%			24.3%		22.1%	
Level 3	Meeting Standard			33.2%			35.2%	35.1%
Level 4		23.1%			31.4%		34.3%	

Performance distributions were also examined separately for males and females. Figure 8.1 shows the percentage of students classified into each performance level for both genders across all grades. Females have a higher percentage of meeting the standard in all grades except grades 4 and 10. The difference between genders is less significant for grades 5 and 7. In Grade 10, a higher passing rate for males is observed.

Figure 8.1 2006 WAAS Performance Distribution by Gender



APPENDIX A. SPECIAL EDUCATION ASSESSMENT ADVISORY PANEL—2006 MEMBERSHIP

First Name	Last Name	Organization	Professional Title
Patricia	Almond PHD	University of Oregon	Researcher, Behavioral Research and Teaching
David	Anderson	OSPI	Special Education Research Analyst
Sherry	Booth	Spokane Public Schools	High School Resource Specialist - Special Education
Becky	Bouchard	Highline School District	Assessment, Eval. And System Accountability
Shawn	Dickson	Aberdeen School District	Special Services Director
Jeanette	Forman	Longview School District	Teacher
Douglas	Gill	OSPI	Special Programs
Sandy	Grummick	OSPI	Program Supervisor, Data Management Services
Bob	Harmon	OSPI	Assistant Superintendent, Special Programs
Jeanne	Hartman	Pearson Educational Measurement	WASL Program Manger
Rick	Hauan	Washington School for the Deaf	Assistant Superintendent
Beth	Hawkins	Pearson Educational Measurement	WASL Program Manager
Marita	Heckart	Puyallup School District	teacher/Self-Contained Program
Eric	Hong	Kent School District	Kentridge High School - Assistant Principal
Mike	Jacobsen	White River School District	Student Support Coordinator
Carol	Johnson	Richland School District	Assistant Director of Special Education
Ann	Koch	Kent School District	Parents
Dick	Koch	Kent School District	Parents
Ed	Koehl	Northshore School District Admin Center	Co-Director for Secondary Sepcial Educational Programs
Judy	Kraft	OSPI	Alternate Assessment Specialist
Hans	Landig	Olympia School District	Executive Director, Student Support
Kristen	Lee	Spokane Public Schools	Special Education Teacher/Intermediate DI Program
YoonSun	Lee	OSPI	Director of Assessment
Colleen	Lines	Washington State School for the Blind	Instructional Resource & Braille Access Centers
Gordon	Linse	Puget Sound Edu. Serv. District 121	Assistant Executive Director, Learning and Teaching
Julie	Moore	Central Kitsap School District	Teacher
Kimberly	Müeller PHD	Tacoma Public Schools	Culminating Project Coordinator
Lucille	Nollette	Bellingham Public Schools	Assistant Director of Special Programs
Molly	O'Conner	OSPI	PR for Assessment Department

Collette	O'Sullivan	OSPI	Alternate Assessment Assistant
Donna K.	Parry	Pearson Educational Measurement	WAAS Project Manager
Gayle	Pauley	OSPI	Director, Title I/LAP and Title V Special Programs
Nina	Potter PHD	Shoreline School District	Assessment Director
Roger	Smith	Sumner School District	Co-Director for Special Services
Deb	Tannasse	Kent School District	Parent & Para educator
Cathy	Taylor	OSPI	Director of Assessment and Innovations
Barbara	Tompkins	Spokane Public Schools	Administrative Assistant / Speech Language Pathologist
Kathie	Weber	Kent School District	Teacher
Joe	Willhoft	OSPI	Assistant Superintendent

APPENDIX B. SPECIAL EDUCATION ASSESSMENT LEADERSHIP TEAM (SEALT)—2006 MEMBERSHIP

The purpose of SEALT is to build capacity for professional learning in the area of assessment of special education students in order to improve student learning. The members of SEALT have been trained to deliver workshops on assessment topics in each region of the state.

FIRSTNAME	LASTNAME	SCH DIST.	SCHOOL
Patricia	Almond PHD		University of Oregon
Gretchen	Baar	Highline SD	Highline SD
Amy	Baldwin	South Kitsap	Hidden Creek
Carol	Banks	Tahoma	Tahoma SD
Linda	Barnes	Shelton	Shelton High School
Niki	Barnes	Kelso	Kelso HS
Candy	Brooks	Shelton	Mt View Elementary
Rayanne	Brown	South Kitsap	Madrona Heights
Sue	Corak	Puyallup	Spinning Elementary
Kelley	Correio	Kennewick School District	Southridge High School
Jessica	Dadisman	Bethel	Orting HS
Christine	Daley	Prosser	Prosser SD
Dalrae	Danilson	Mount Vernon	Washington Elementary
Edward	Durkin	Taholah	Taholah
Holly	Eigenbrood	Shoreline	Shoreline SD
Cherry	Fairbanks	Granger	Granger SD
Karen	Fisher	Methow Valley	Methow Valley Elementary
Jeanette	Forman	Longview	Cascade MS
Heidi	Freeman	Evergreen	Covington Middle School
Faye	Fuchs	ESD 105	ESD 105
Glenn	Green	Central Valley	Learning and Teaching Center
Carol	Hall		
Kathleen	Harvester	East Valley SD #90	East Valley HS
Donnita	Hawkins	OSPI	Reading Dept.
Marita	Heckert	Puyallup	Ballou Junior High
Valerie	Hooton	Arlington	Arlington HS
Carissa	Jangard	Anacortes	Whitney Elementary
Kathy	Kipp	Colville #115	Aster Elementary - Colville J H
Judy	Kraft	OSPI	OSPI
Jeannene	London	Kennewick	Highlands MS
Mo	Lyden	ESD 101	ESD 101
Alan	Mahler	CVSD 356	University High School
Betty	McKinney	Ellensburg	Morgan Middle School
Tonya	Middling	OSPI	Special Education dept.
Marcia	Mongrain	Kelso	Cowlitz County Youth Services
Maggie	Morelan	Toppenish	ESD 105
Carolyn	Munoz	Camas	Skyridge MS
Molly	O'Conner	OSPI	OSPI

Collette	O'Sullivan	OSPI	OSPI
Donna	Perry	Pearson	Pearson Educational Measurement
Kerry	Salaz	LaConner	LaConner Elementary
Audrey	Seaberg	Quincy	Quincy SD-Support Services
Kay	Sidwell	Moses Lake	Moses Lake HS
Carrie	Stokes	NCESD	North Central Educational Schoold District
Robin	Taylor	Brewster	Brewster SD
Catherine	Taylor	OSPI	OSPI
David	Toston		
S. David	Tudor	Washougal	Jemtegaard
Becky	Turnbull	ESD 113	Lewis County Special Education Coop
Laura	Vanderhoof	Tenino	Parkside Elementary
Joe	Willhoft	OSPI	OSPI
Michele	Wollert	WA State	School for the Blind

***APPENDIX C. SPECIAL EDUCATION CONTENT WORKGROUPS—2006
MEMBERSHIP***

Special Educator/General Educator/Curriculum Specialist	Content	Special Educator/General Educator/ Curriculum Specialist	Content
Helen Adams	Math	Luanne Clark	Writing
Janet Bliss	Math	Scott Cooley	Math
Debra Blodget	Math	Diane Ganey	Math
Rebecca Butters	Math	Beth Harrington	Reading
Ann Cadwallader	Math	Jeri Lynn Howard	Math
Eileen Clausen	Reading	Kelly Jacobsen	Math
Randy Cloke	Math	Jo Ellen Kain	Writing
Nicole Connor	Reading	Katie Lloyd	Math
Sue Corak	Reading	Diana McFaul	Reading
Dalrae Danilson	Reading	Tom Pollino	Reading
Darci Downs	Math	Amie Ojerio	Science
Martha DuMerton	Writing	Anne Ryan	Science
Jeanette Forman	Writing	Carolyn Brown	Science
Runell Anne Galina	Reading	Clinton C. Winn	Science
Laura Hendrickson	Writing	Del Guenther	Science
Mickey Johnston	Writing	Emily Garlich	Science
Juanita Kamphuis	Math	Georgia Boatman	Science
Dawnie Kelly	Writing	Irma Lange	Science
Rhonda Larson	Math	Jason Maioho	Science
Sandi Laurie	Reading	Jeanette Forman	Science
Georgina Lynch	Reading	Karen Plum	Science
Chuck Moore	Math	Kathy Colombo	Science
Maggie Morelan	Math	Katie Thieme	Science
Ellen Nickerson	Math	Katy Taylor	Science
Billie Parke	Reading	Kim Clark	Science

Special Educator/General Educator/Curriculum Specialist	Content	Special Educator/General Educator/ Curriculum Specialist	Content
Lisa Parvey	Writing	Laura Vanderhoof	Science
Jennifer Presley	Writing	Lauri Anderson	Science
Paulette Rich	Writing	Lisa Chi	Science
Stacey Rockey	Writing	Marcia Gwynne	Science
Elizabeth Scott	Writing	Marj Hendricks	Science
Carrie Stokes	Reading	Melinda Poindexter	Science
Laura Vanderhoof	Reading	Monica Johnston	Science
Sandra Watson	Reading	Morgan Jones	Science
Amy White	Reading	Patrice (Patti) Adkinsson	Science
Terry Frentrop	Science	Paul Baldwin	Science
Sarah Pooler	Science	Sarah Mitchell	Science
Rob Stagg	Science		

APPENDIX D. PARTICIPATION GUIDELINES FOR THE 2006 WAAS

How Should Students in Special Education Programs Participate in State Assessments?

Students must participate in state assessments in all content areas scheduled for testing based on the student's grade level. In cases where the student is enrolled in a non-graded program, the student is assessed in the same content areas at the equivalent chronological age as a student in the grades assessed. Students with disabilities are expected to take the Washington Assessment of Student Learning (WASL) tests whenever possible, with or without necessary accommodations, unless the Individualized Education Program (IEP) team determines that the student is unable to participate on the WASL in one or more content area. In this case, the IEP team may select the Washington Alternate Assessment System (WAAS) portfolio assessment. If the student is unable to take the WASL, even with accommodations, then the student must participate by taking the WAAS portfolio assessment.

The student's IEP team must meet to determine:

- how the student will participate in the WASL;
- which testing accommodations, if any, should be provided; or
- if the student requires an alternate assessment.

The IEP decisions regarding testing accommodations or why the WASL is not appropriate and how the student will be alternately assessed must be documented in the student's current IEP.

Guidelines for IEP Team Decision-Making in Regard to Participation in the State Assessment System Guidelines are intended to inform the decisions of IEP teams regarding WASL participation—based on the needs of each student—resulting from his or her disability. IEP teams have the authority to determine the manner in which a student with a disability will participate in WASL assessments. These decisions must be made separately in each content area in which the student is scheduled for statewide assessment the following spring. For example, a student may take the standard WASL test with or without accommodations in certain subjects but may require alternate assessment in other subjects.

Following are three case studies designed to provide more guidance for IEP teams in determining which assessment option is appropriate for students in special education programs. In each case study, characteristics of the student's instructional program and participation in classroom assessments are described. The recommendations for participation in the assessment options are also included in the case studies.

Case Study 1

Instructional Program: The student is engaged in an instructional program guided by the EALRs in this content area and is working on benchmarks at or near grade-level expectations.

Classroom Assessment: The student is generally able to take a paper-and-pencil test under routine conditions or with testing accommodations.

State Assessment Option: The student should take the standard WASL in this content area. The student may need testing accommodation(s) that are modeled on instructional accommodation(s) used in the student's educational program. Refer to the Guidelines for Participation and Testing Accommodations for Special Populations in State Assessment Programs.

Case Study 2

Instructional Program: The student is engaged in an instructional program guided by the EALRs in this content area but is working on EALRs that have been modified to reflect below grade-level expectations for performance due to the nature of the student's disability.

Classroom Assessment: The student is generally able to take a paper-and-pencil test under routine conditions or with one or more testing accommodations.

State Assessment Option: The student should take the standard WASL in this content area, with any necessary testing accommodation(s) that are modeled on instructional accommodation(s) used in the student's educational program. Refer to the Guidelines for Participation and Testing Accommodations for Special Populations in State Assessment Programs.

Case Study 3

Instructional Program: The student is engaged in an instructional program guided by the EALRs in this content area substantially below any grade-level expectations and may be focused on EALR

Extensions due to the nature and severity of the student's disability or disabilities and, such that:

- these disabilities severely limit the student's involvement in the EALRs even with program modifications and adaptations; and
- the student requires intensive, individualized instruction in multiple settings in order to acquire knowledge and to accomplish the transfer and generalization of skills in this content area to school, work, home, and community.

Classroom Assessment: The student is generally unable to demonstrate knowledge on a paper-and-pencil test, even with accommodations.

State Assessment Option: The student should participate through the WAAS portfolio in this content area. Participation in alternate assessments is intended for a very small number of students with significant disabilities who are unable to participate on the WASL, even with accommodations.

APPENDIX E. WAAS PORTFOLIO SCORING 2006 PROCEDURAL STEPS

Scorer Procedures

- The first scorer performs the following tasks:
 1. Go to checklist at assigned Binder Checkout Station and select a district and binder number to score. The scorer **cannot** select a district to score if that district is represented at your table.
 2. Write employee number on the checklist under first scorer next to the binder number.
 3. Located inside the binder are the scoring monitors. The scorer will pull and grid his/her employee number on scoring monitor 1 before scoring begins. Verify that the bar-coded binder number matches the scoring monitor number.
 4. Score the portfolio (use scorer note sheet if desired).
 - Read the student profile for contextual information
 - Scan the entire portfolio
 - Look at skill on the entry cover sheet
 - Look for data/evidence of progress over time on stated skill in entry
 - Check for criteria for meeting the goal
 - Compare evidence to criteria to determine if student met/exceeded the goal
 - Secondary evidence may provide more information on the skill performance
 - Determine score using scoring criteria
 - Score Part 1 in Reading, Writing, Mathematics, and/or Science
 - Look for generalization of skills across the portfolio
 - Score modifications/adaptations using scoring criteria
 - Score settings/contexts using scoring criteria
 - Score social relationships using scoring criteria
 - Score self-determination using scoring criteria
 5. Score and grid scores on monitor 1. Write employee number and grid monitors 6–9.
 6. Return binder with WAAS demographic sheet to table 1. Initial next to the binder number on the checklist that the binder has been returned.
 7. Return Scoring Monitor 1 and 6–9 to trays at Binder Checkout Station.

- The second scorer performs the following tasks:
 1. Go to the checklist, select a district and binder number to score. The scorer **cannot** select a district to score if that district is represented at your table.
 2. Write employee number on the checklist under second scored next to the binder number. Second scoring can only occur if binders have been scored previously.
 3. Located inside the binder are the scoring monitors. The scorer will grid his/her employee number on scoring monitor 2 before scoring begins.
 4. Score the portfolio (use the scorer note sheet if desired). See step #4 above for details.
 5. Score and grid scores on monitor 2
 6. Return the binder to table 2 and initial next to the binder number on the checklist that the binder has been returned.
 8. Return Scoring Monitor 2 to Binder Checkout Station.

Table Leader Procedures

1. Backread at least one scored portfolio for each scorer each day. Your score will be the second read.
2. When not backreading a portfolio for someone at your table, go to a Checkout Station (different from your assigned station) to select a portfolio to second read. Write your employee number as second reader on the checklist.
3. Score portfolios using the procedure above. Check to make sure the scorer employee number and scoring monitors 1 and 6–9 are done (when backreading) or turned in (during second reads).
4. Note and retrain scorers at your table if they are discrepant with your scores.
5. Complete Scoring Monitor 2.
6. Return the scored portfolio to the correct Binder Checkout Station and initial that the binder was returned. Put monitor 2 in the tray. If it was a backread for someone at your table, sign in and initial as second scorer.
7. Assist scorers when they have questions.

APPENDIX F. WAAS PORTFOLIO SCORING SUMMARY SHEET—2006

Part I: Progress on IEP Skills (Progress on targeted skills scored separately in each content area)

	1	2	3	4
Progress on IEP Skill linked to EALRs	No or little progress on targeted skills linked to the EALRs and GLEs in the portfolio entry.	Clear progress on targeted skills linked to the EALRs and GLEs in the portfolio entry.	Attains goal for targeted skills linked to the EALRs and GLEs in the portfolio entry.	Exceeds goal for targeted skills linked to the EALRs and GLEs in the portfolio entry.
CONTENT AREA				PART I SCORE
Reading (Grades 3–8, H.S.)				
Writing (Grades 4, 7, H.S.)				
Mathematics (Grades 3–8, H.S.)				
Science (Grades 5, 8, H.S.)				

Part II: Student Generalization of Skills (These dimensions are scored across portfolio)

Dimension	1	2	3	4	Scorer Use Only
Modifications and Adaptations	No evidence that the student uses supports, modifications, adaptations, or assistive technology in portfolio entry.	Some evidence that the student appropriately uses supports, modifications, adaptations, or assistive technology in portfolio entry.	Evidence that the student appropriately uses supports, modifications, adaptations, or assistive technology throughout the portfolio entry.	Evidence that the student appropriately uses supports, modifications, adaptations, or assistive technology and that some of the support is natural in portfolio entry.	
Settings and Contexts	Student performs targeted skill in a limited number of settings, or use of targeted skill is unclear in portfolio entry.	Student performs targeted skill in some settings or contexts in portfolio entry.	Student performs targeted skill in a variety of settings or contexts in portfolio entry.	Student performs targeted skill in an extensive variety of settings or contexts in portfolio entry.	
Social Relationships	The student has no social interactions during standards-based instructional activities with others, with and without disabilities, in the portfolio entry	The student has some social interactions during standards-based instructional activities with peers, with or without disabilities, in the portfolio entry.	The student has sustained social interactions during standards-based instructional activities with peers, with or without disabilities, in the portfolio entry.	The student has varied, sustained social interactions during standards-based instructional activities with peers, with and without disabilities, in the portfolio entry	
Self-Determination	The student does not exhibit self-determination skills within activities in the portfolio entry.	The student monitors own learning within activities in the portfolio entry.	The student monitors and evaluates own learning within activities in the portfolio entry.	The student monitors, evaluates, and sets goals for own learning within activities in the portfolio entry.	

CONTENT AREA	PART I SCORE	PART II TOTAL	TOTAL SCORE IN CONTENT AREA
Reading (Grades 3–8, H.S.)			
Writing (Grades 4, 7, H.S.)			
Mathematics (Grades 3–8, H.S.)			
Science (Grades 5, 8, H.S.)			

APPENDIX G. WAAS PORTFOLIO ACADEMIC ACHIEVEMENT STANDARD DESCRIPTIONS

The academic achievement standards for students with significant disabilities who are participating in the Washington Alternate Assessment System (WAAS) portfolio are significantly different from the standards for students who participate in the Washington Assessment of Student Learning (WASL). The WAAS portfolio is based on the Essential Academic Learning Requirements (EALR) Extensions, which allow the student to participate and progress in the general curriculum. Because the WAAS portfolio is based on the student's Individualized Education Program (IEP) goals in relation to the EALR Extensions, the specific assessment targets selected for the student may be the same for many content areas but may be different for any other student. Additionally, these students have educational goals that may remain the same throughout their educational careers. Therefore, the following academic achievement standard descriptors apply for all grades and content areas.

Level 1—Students performing at this level will be making little or no progress toward the goal for the targeted IEP skills linked to the EALRs. The student is unable to generalize the use of these targeted skills, using modifications and adaptations in any settings or contexts. The student cannot make choices in planning, monitoring, or evaluating own performances. The student has no or limited social interactions with others during educational activities.

Level 2—Students performing at this level will be making some progress toward the goal for the targeted IEP skills linked to the EALRs. The student is able to generalize the use of these targeted skills in some ways. The student may appropriately use modifications and adaptations in some settings and contexts or make choices in planning, monitoring, or evaluating own performances. The student may have some social interactions with others during educational activities. The student is not able to generalize the targeted IEP skills in all these ways.

Level 3—Students performing at this level will be making clear progress or attaining the goal for the targeted IEP skills linked to the EALRs. The student is able to generalize the use of these targeted skills, appropriately using modifications and adaptations in a variety of settings and contexts while making choices in planning, monitoring, or evaluating own performances. The student sustains some social interactions with others during educational activities.

Level 4—Students performing at this level will be attaining or exceeding the goal for the targeted IEP skills linked to the EALRs. The student is able to generalize the use of these targeted skills, appropriately using natural supports, modifications, or adaptations in an extensive variety of settings or contexts while consistently making choices in planning, monitoring, or evaluating own performances. The student has sustained, varied social interactions with others during educational activities.