THE WASL: A CRITICAL REPORT TO INTERESTED CITIZENS OF THE STATE OF WASHINGTON

Dr. Donald C. Orlich

March 15, 2005

Executive Summary

Conclusions. This report is an analysis of the 2004 Grade 5 Science WASL and the Grade 7 Mathematics WASL using criteria from developmental psychology and the scales of the National Assessment of Educational Progress (NAEP). Inferences from this study may be applicable to the entire battery of WASL assessments.

1. The Grade 5 Science WASL exceeds the intellectual level of the vast majority of grade 5 children and appears to be an 8th grade examination.
2. While not specifically examined, English language learners will find this assessment to be virtually impossible to pass due to needed vocabulary skills.
3. The Grade Level Expectations (GLE's) for Grade 5 science are developmentally inappropriate. The GLE's drive the WASL; thus the test is developmentally inappropriate.
4. The 7th Grade Math WASL is in all reality a 9th grade test.
5. Test items do not progress from relatively easy to more difficult. They simply appear with no logical sequence. Standardized tests begin with easy items and move to more difficult ones.
6. A total of 9 math concepts are tested. Yet, 185 math General Level Expectations are listed for Grade 7.
7. Reading and writing are most critical for student success. One could hypothesize a very high correlation between these two skills and success in the Science and Mathematics WASL.
8. Reviewing the GLE's for grade 7 and 10 reveals parallel entries. That is, the grade 7 GLEs are almost identical in many cases to those of grade 10.

Policy Implications. There are instructional and policy implications associated with the findings and conclusions of this analysis.

First, if the WASL tests are advanced beyond the mental cognition of grade 5 and 7 pupils, then for most children failure will be the ultimate end, regardless of instructional techniques used.

Second, what psychological impact will failing an inappropriate science and math WASL have on students and their ultimate attitudes towards science and math, and schooling in general?

Third, one may predict litigation by concerned parents and child advocacy groups against the State of Washington.

Fourth, scoring errors have been found nationally in virtually all mandatory high-stakes tests. These have led to class action law suits. For example, the state of Minnesota paid out approximately $12 million to students and/or their parents due to scoring errors.

Fifth, the legislature is approaching fiscal irresponsibility or is not practicing fiscal accountability by continuing to fund the exorbitant WASL. With the State of Washington viewing at least a $2.2 Billion budget short fall, the massive $200,000,000 OSPI budget for school reform must be challenged.

Sixth, the legislature should commission an outside research organization to verify or refute this study.

INTRODUCTION

Educational reform in Washington State has in reality been reduced to "Doing the WASL," *Washington Assessment of Student Learning*. This high-stakes test in mathematics, reading and writing is administered each spring to 4th, 7th and 10th graders. Science is mandatory in grades 5, 8 and 10. Reading and Math WASLs are being developed for grades 3, 5, 6 and 8 to fulfill federal requirements agreed to by the Office of Superintendent of Public Instruction. School children and educators take the WASL very seriously. There is much propaganda that we need to push students to their limits by raising the bar. However, this assumption is based on a flawed premise, as will be demonstrated with empirical data.

While much has been publicized about the WASL no one has analyzed the actual tests using published and long-accepted criteria that have stood the test of time. The focus of this report is on the 2004 Grade 5 Science WASL (see Part I) and the Grade 7 Mathematics WASL (see Part II). Inferences from this study may be applicable to the entire battery of WASL assessments.

Establishing the Limits to Student Achievement

To initiate my premise that there is a limit to the quantity and quality of student achievement, albeit not fixed, the Developmental Perspective will be used. This approach is associated with Jean Piaget (1969). His model assumes that humans evolve intellectually in various overlapping stages. Piaget describes four stages or periods of development—the *sensorimotor stage*, from birth to two years; the *preoperational stage*, from two to eight years; the *concrete operational stage*, from eight to eleven years; and the *formal stage*, from eleven to fifteen years and up.

The last stage is what schools attempt to reach in what we generally call thinking and analyzing. However, the majority of students in middle and high school are still in the concrete developmental stages. The listing below summarizes the developmental stages and adds the behavioral model of cognitive development, known as “Bloom’s Taxonomy” (Bloom et al. 1956). The latter approximates the National Assessment of Progress Levels (NAEP).

Epstein/Piaget Developmental Levels

1. *Entry concrete,* e.g., orders a series but would not observe relationships.
2. *Advanced concrete,* e.g., identifies one variable that affects results.
3. *Entry formal*, e.g., seeks “why” some phenomenon takes place and identifies causes.
4. *Middle formal*, e.g., interprets higher order graphical relationships.

Bloom’s Taxonomy Levels

1. *Knowledge*, e.g., recalls or recognizes information.
2. *Comprehension,* e.g., states examples in own words.
3. *Application*, e.g., uses information to solve problems.
4. *Analysis*, e.g., identifies issues or implications, and isolates component parts.
5. *Synthesis*, e.g., creates new forms or identifies abstract relationships.
6. *Evaluation*, e.g., judges via criteria.

Table 1 provides the relative percentages of students at Piaget’s stages of development as synthesized by Herman T. Epstein (see 2002), a world authority on the subject.Table 2 illustrates what cognitive tasks children can do at various levels assembled by two international authorities, Michael Shayer and Philip Adey (1981).These data form the basis of my interpretation of Tables 3-6, which present published data from the NAEP ages 9, 13 and 17 in science, mathematics, reading; and for grades four, eight and eleven in writing.

TABLE 1. PERCENTAGE OF STUDENTS AT PIAGET'S COGNITIVE LEVE*LS*

Age	*Grade*	Intuition	Entry Concrete (a)	Advanced Concrete (b)	Entry Formal (a)	Middle Formal (b)	Ref.
5.5	P	78	22				J
6	K	68	27	5			A
7	1	35	55	10			A,W
8	2	25	55	20			A
9	3	15	55	30			A
10	4	12	52	35	1		S
11	5	6	49	40	5		S
12	6-7	5	32	51	12		S
13	7-8	2	34	44	14	6	S
14	8-9	1	32	43	15	9	S
15	9-10	1	15	53	18	13	S
16	10-11	1	13	50	17	19	S
16-17	11-12	3	19	47	19	12	R
17-18	12	1	15	50	15	19	R
Adult	---	20	22	26	17	15	R

Table 1. Notes and References

1. Level (a) in each category is composed of children who have just begun to manifest one or two of that level's reasoning schemes, while level (b) refers to children manifesting a half dozen or more reasoning schemes.

2. Table derived by Herman T. Epstein, personal communication, June 8, 1999. See also: Herman T. Epstein, "Biopsychological Aspects of Memory and Education." In S. P. Shohov, Editor, *Advances in Psychology Research, Volume 11.* New York: Nova Science Publisher, Inc. , 2002, pp. 181-186

J Smedslund, J. (1964). *Concrete Reasoning: A Study of Intellectual Development.* Lafayette, IN: Child Development Publications of the Society for Research in Child Development.

A Arlin, P. Personal Communication with H. T. Epstein.

W Wei, T. D., et al. (1971). "Piaget's Concept of Classification: A Comparative Study of Socially Disadvantaged and Middle-Class Young Children." *Child Development* (42): 919-927.

R Renner, J. W., Stafford, D. G., Lawson, A. E., McKinnon, J. W., Friot, F. E. and Kellogg, D. H. (1976). *Research, Teaching and Learning With the Piaget Model*. Norman: University of Oklahoma Press.

S Shayer, M. and Adey, P. (1981). *Towards a Science of Science Teaching.* London: Heinemann.

TABLE 2. SELECTED CONCEPTS WITH PIAGETIAN DESCRIPTORS ILLUSTRATING CONCRETE TO FORMAL DEVELOPMENT OF A CHILD'S INTERACTION WITH THE WORLD

Topic	Early Concrete	Late Concrete	Early Formal	Late Formal
Investigative Style	Unaided style does not produce models	Can serially order and classify objects	Is confused, needs an interpretive model	Generates and checks possible explanations
Relationships	Can order a series but cannot make summa-rization	Readily uses the notion of reversibility	Can begin to use two independent variables	Reflects on reciprocal relationship between variables
Use of Models	Simple comparisons-- one to one corre-spondence	Simple models, e.g., gear-box, skeleton	Deductive com-parisons and models are taken as being true	Searches for explanatory model, uses proportional thinking
Categorizations	Objects are classified by one criterion- color, size	Partially orders and classifies hierarchi-cally	Generalizes to impose meaning over wide range of phenomena	Abstract ideas gener-ated-- searches for underlying associa-tions
Proportionality	Needs whole sets to double or halve	Makes inferences from constant ratios and with whole num-bers only	Makes inferences on ratio variables- Density = Mass/Volume	Knows direct and inverse relationship ratios
Mathematical Opera-tions	Number is distin-guished from size or shape	Works single opera-tions but needs closure	Generalizes by concrete examples and accepts lack of closure	Conceives of a vari-able properly
Probabilistic Thinking	No notions of prob-ability	Given equal number of objects knows there is 50/50 chance of one being drawn	Given set of objects can express chances in simple fractions	

Source: Michael Shayer and Philip Adey. *Towards a Science of Science Teaching: Cognitive Development and Curriculum Demand, 1981.* London: Heinemann. Abstracted from Table 8.1, pp.72-78.

Please note that Shayer and Adey did much of their work with "clever" children, most having IQ's of 160 and up.

Examining the NAEP Data with Developmental Criteria

Table 1 illustrates the relative percentages of school-aged children and their cognitive levels. Note that until grade 4 (ages 9 or 10) that 100 to 99 percent of children, respectively, are yet in the concrete or intuitive levels of cognition. Examine Tables 2, 3, 4, 5 and 6. Observe how from data in Table 1 one could predict that zero percent of the nine-year olds would be able to answer questions on the NAEP 350 Level! This is evidence that can only be interpreted that over a 20-year period of time, no nine-year olds in the NAEP national samples are capable of answering the higher level thinking items on the NAEP tests. One can equate the NAEP 350 Level with "Bloom's Taxonomy" Levels of synthesis and evaluation or the so-called "higher-order of thinking" domains.

Conversely, observe the gradual decrease in the sampled fourth grader percentages correct by moving from Levels 150 to 250. At NAEP Level 150, the percentages range from 91 to 99 percent. No question, these are concrete cognition problems, along with NAEP Level 250. One would predict the downward scores from the Table 1 descriptions of the cognitive levels. It appears that the critical level for fourth graders is NAEP Level 250 or the equivalent of Bloom's Application Level.

Observe parallel patterns for 13 and 17-year olds. These American youth do brilliantly at NAEP Levels 150 and 200, as one can predict from the tabular set.

TABLE 3. PERCENTAGES OF STUDENTS PERFORMING AT OR ABOVE SCIENCE PERFORMANCE LEVELS, AGES 9, 13 AND 17, 1977 AND 1996.

		AGE 9		*AGE 13*		*AGE 17*	
Level		**Percent in 1977**	**Percent in 1996**	**Percent in 1977**	**Percent in 1996**	**Percent in 1977**	**Percent in 1996**
350	Can infer relationships and draw conclusions using detailed scientific knowledge.	0	0	1	0	9	11
300	Has some detailed scientific knowledge and can evaluate the appropriateness of scientific procedures.	3	4	11	12	42	48*
250	Understands and applies general information from the life and physical sciences.	26	32*	49	58*	82	84
200	Understands some simple principles and has some knowledge, for example, about plants and animals.	68	76*	86	92*	97	98
150	Knows everyday science facts	94	97*	99	100*	100	100

* Indicates that the percentage in 1996 is significantly different from that in 1977.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP). *Report in Brief, NAEP 1996 Trends in Academic Prog??ress.* Revised 1998. NCES 98-530, Table 1, p. 9.

TABLE 4. PERCENTAGES OF STUDENTS PERFORMING AT OR ABOVE MATHEMATICS PERFORMANCE LEVELS, AGES 9, 13 AND 17, 1978 AND 1996.

Level		*AGE 9*		*AGE 13*		*AGE 17*	
		Percent in 1978	**Percent in 1996**	**Percent in 1978**	**Percent in 1996**	**Percent in 1978**	**Percent in 1996**
350	Can solve multi-step problems and use beginning algebra.	0	0	1	1	7	7
300	Can compute with decimals, fractions and percents; recognize geometric figures; solve simple equations; and use moderately complex reasoning.	1	2*	18	21	52	60*
250	Can add, subtract, multiply and divide using whole numbers and solve one-step problems.	20	30*	65	79*	92	97*
200	Can add and subtract two-digit numbers and recognize relationships among coins.	70	82*	95	99*	100	100
150	Knows some addition and subtraction facts.	97	99*	100	100	100	100

* Indicates that the percentage in 1996 is significantly different from that in 1978.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP). *Report in Brief, NAEP 1996 Trends in Academic Prog??ress.* Revised 1998. NCES 98-530, Table 2, p. 10.

TABLE 5. PERCENTAGES OF STUDENTS PERFORMING AT OR ABOVE READING PERFORMANCE LEVELS, AGES 9, 13 AND 17, 1971 AND 1996.

Level		*AGE 9*		*AGE 13*		*AGE 17*	
		Percent in 1971	**Percent in 1996**	**Percent in 1971**	**Percent in 1996**	**Percent in 1971**	**Percent in 1996**
350	Can synthesize and learn from specialized reading materials.	0	0	0	1*	7	6
300	Can find, understand, summarize and explain relatively complicated information.	1	1	10	14*	39	39
250	Can search for specific information, interrelate ideas and make generalizations.	16	18*	58	61*	79	81*
200	Can comprehend specific or sequentially related information.	59	64*	93	93	96	97*
150	Can carry out simple, discrete reading tasks.	91	93*	100	100	100	100

* Indicates that the percentage in 1996 is significantly different from that in 1971.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP). *Report in Brief, NAEP 1996 Trends in Academic Prog??ress.* Revised 1998. NCES 98-530, Table 3, p. 11.

TABLE 6. PERCENTAGES OF STUDENTS PERFORMING AT OR ABOVE WRITING PERFORMANCE LEVELS, GRADES 4, 8 AND 11, 1984 AND 1996.

Level		***GRADE 4***		***GRADE 8***		***GRADE 11***	
		Percent in 1984	**Percent in 1996**	**Percent in 1984**	**Percent in 1996**	**Percent in 1984**	**Percent in 1996**
350	Can write effective responses containing details and discussion.	0	0	0	1	2	2
300	Can write complete responses containing sufficient information.	1	1	13	16	39	31*
250	Can begin to write focused and clear responses to tasks.	10	13	72	66*	89	83*
200	Can write partial or vague responses to tasks.	54	59	98	96*	100	99
150	Can respond to tasks in abbreviated, disjointed or unclear ways.	93	93	100	100	100	100

* Indicates that the percentage in 1996 is significantly different from that in 1984.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP). *Report in Brief, NAEP 1996 Trends in Academic Prog??ress.* Revised 1998. NCES 98-530, Table 4, p. 12.

PART I

Analyzing the Science WASL

The previous and rather lengthy introduction shows the analytic tools used to determine the intellectual appropriateness of the WASL tests analyzed in this report.

The Grade 5, 2004, Science Washington Assessment of Student Learning (WASL) is constructed primarily of six written "scenarios." This design requires all children to be at least above average readers. To be successful, it is imperative that all children have extensive graphing and tabular interpreting experiences. The test has no apparent simple to complex arrangement of test items. The 38 test items begin with a most difficult scientific process—*experimental design*. Any child who has not mastered the processes of analyzing and designing single-variable experiments will fail because 50% of the entire WASL test relates to experimental design.

Those test items not included in the scenarios simply "pop-up." That is, placement of the items in the test booklet appears to be at random. Generally, standardized tests begin with rather easy items and then progress to more difficult ones. Such test construction provides for student self-confidence.

Of the 38 test items 12 (32%) are multiple choice. The remaining 26 items (68%) of the WASL require extensive written responses (and one might add very subjective evaluation—the use of scoring "rubrics" notwithstanding). Reading and writing are the primary skills being assessed, because at most, only seven or eight science concepts are included. As noted, the "pop-up" items relate to astronomy, human skin, roots, tools and volcanoes. These questions show little relationship to a coherent body of science content. The absurdity of these "pop-ups" is demonstrated in multiple choice question number 35, asking children to determine the difference between a tree and a swimming mammal! Such questions make a mockery of scientific literacy.

A series of five questions relate to *Mechanical Advantage*. This concept is algebraically derived and is totally inappropriate for Grade 5 youngsters (review Table 4).

Table 7 provides a brief content analysis of the broad science foci or concepts being tested on the Grade 5 WASL.

All the WASL scenarios require advanced concrete or formal cognition. Thus, a vast majority of children at grade 5 will have difficulty understanding the test. Re-examine Table 3, levels 300 and 350 for children in grades 4 and 8. No sampled child in America at ages 9 and 13 in grades 4 and 8 could answer questions at the level 350, that is, "Can infer relationships and draw conclusions using detailed scientific knowledge." Scan Table 1 and observe that only 45% of the children at this age/grade level are in the advanced or early formal stage. Further, children do not reach these stages simultaneously: Some are ahead and others are behind the stage.

TABLE 7. CONTENT ANALYSIS OF GRADE 5 SCIENCE, 2004 WASHINGTON ASSESSMENT OF STUDENT LEARNING (WASL)

Focus or Concept	No. of Items	Percent of Total
1. Experimental Design (Light, Bubbles, Erosion)	19	50%
2. General Biological Sciences	6	16%
3. Mechanical Advantage	5	13%
4. Scientific Equipment/Tools	3	8%
5. Energy	2	5.5%
6. Pollution	2	5.5%
7. Volcanoes	1	3%
TOTALS	**38**	**100%**

At NAEP Level 300, the testing focus is "Has some detailed scientific knowledge and can evaluate the appropriateness of scientific procedures." That means students can master experimental design, the major focus of the Grade 5 WASL. Note that only 4% of the sampled children ages 9 (Grade 4) and only 12% of children age 13 (Grade 8) could respond correctly to those NAEP Level 300 test items. Again, this single process concept accounts for 50% of the Science WASL.

The NAEP data substantiate and validate the Epstein/Piaget data almost perfectly. These data give us a predictive validity showing a very high failing rate (not meeting standard) for grade 5 children taking the Science WASL.

In 2001, my predictions of students who would fail to meet standard on the Science WASL were published based on a detailed analysis only of the then called "Essential Academic Learning Requirements," now relabeled "Grade Level Expectations." For Grade 5, Science WASL, I predicted a most pessimistic failure rate of 63% - 67%. In Spring of 2004, 71.8% failed. For grade 8, my prediction was 60% - 64% fail. In Spring 2004, 60.6% failed. For grade 10, my prediction was that 60% to 64% would fail. In Spring 2004, 67.8% failed. My predictions, based on the (so-called) state science standards, were all within 4% of being perfect. This can only happen when the tests are developmentally inappropriate. It certainly appears that the OSPI and WASL writers are confusing "rigor" with "developmentally advanced." Or they are confusing both.

Conclusions About Grade 5 Science WASL

Based on the analysis shown in this paper, there are five major conclusions.

1. The Grade 5 Science WASL exceeds the intellectual level of the vast majority of grade 5 children.

2. Reading and writing are most critical for student success. One could hypothesize a very high correlation between these two skills and success in the Science WASL.

3. While not specifically examined, English language learners will find this assessment to be virtually impossible to pass due to needed vocabulary skills.

4. The Grade Level Expectations (GLE's) for Grade 5 science are developmentally inappropriate. The GLE's drive the WASL; the test is therefore developmentally inappropriate.

5. No amount of tinkering with this assessment can make it appropriate for grade 5. This is a grade 8 assessment (please refer to Tables 1, 2 and 3.)

PART II

Analysis of Grade 7, 2004 Mathematics WASL

The Grade 7, 2004 Mathematics WASL is, like the Science WASL, keyed to the Essential Academic Learning Requirements (EALRs) and now labeled "Grade Level Expectations." First, let us quickly critique the original EALRs for math.

A Bag of Tools or Tools that Fool?

There are five basic EALRs for Mathematics as is noted in the box below.

1. The student understands and applies the concepts and procedures of mathematics.
2. The student uses mathematics to define and solve problems.
3. The student uses mathematical reasoning.

4. The student communicates knowledge and understanding in both everyday and mathematical language.
5. The student understands how mathematical ideas connect within mathematics, other subject areas, and real-life situations.

These five broad standards are then subdivided into 18 subcategories, e.g., standard "1.3 Understand and apply concepts and procedures from geometric sense—properties and relationships and locations and transformations." These substandards are further subdivided into 66 "Benchmarks" for grades 4, 7 and 10. That care and critical analyses were disregarded is evidenced by the following examples from the benchmark on Estimation.

"Grade 4: Use estimation to predict computation results and to determine the reasonableness of answers, *for example, estimating a grocery bill.*

Grade 7: Use estimation to predict computation results and to determine the reasonableness of answers involving nonnegative rational numbers, *for example, estimating a tip.*

Grade 10: Use estimation to predict computation results and to determine the reasonableness of answers involving real number, *for example, estimating."*

Please re-read that list. The same standard lead-in is applied to fourth, seventh and tenth graders. Are we missing something here? There are between twelve and fourteen similar sets of identical standards for the three grades. Did anyone mention an editing problem? This is not applying the long used "spiral curriculum model" in which concepts are introduced at one grade level and then expanded and elaborated in others. No, the EALRs were rushed into print for pure political reasons and show total disregard for developmental appropriateness or logic.

My favorite "tool that fools" EALR for fourth graders is from Standard 1.5: **"Understand and apply concepts and procedures from algebraic sense. Benchmark 1, Grade 4, To meet this standard the student will: Write a rule for a pattern based on a single arithmetic operation between terms such as a function machine."**

A "function machine"? Where can I buy one: Is it like a laptop? Citizens, this is what politically motivated individuals in Washington State think kids need to be successful in the 21st Century. (Yours truly has been very successful for the 20th and now the 21st Century and I don't have a clue what a "function machine" is! Do you?) Keep in mind that by the year 2008, a student will be denied a high school diploma if he or she does not pass the WASL in the 10th grade.

The Revised 2005 Math Grade Level Expectations

If there were problems with the various iterations of the math EALRs, they are certainly exaggerated with the publication of the 2005 Grade Level Expectations (GLEs). Very specific student expectations are listed. The entire set of GLEs is nothing less than a series of mechanistic performance objectives wherein a specific behavior is to be elicited from the students. (Children are treated as if they were programmable machines.) The teacher can then observe whether learning has taken place or not. That there appears to be editing problems once again is shown in just one example selected at random from the set.

Grade 7, GLE 1.5.2. "Write a story about a situation that represents a given linear equation, expression or graph."

Grade 10, GLE 1.5.2. "Write a story that represents a given linear equation or expression."

Educators must challenge the appropriateness of identical GLEs for grade 7 and 10. The example cited is not one-of-a-kind. A careful examination of the entire list of GLEs shows that in many cases grade 7 and grade 10 students are required to do the same mathematical operation, reason logically, solve problems, communicate understanding or make connections from the classroom to the outside world.

I will not argue the merits of the philosophy or psychology being touted by the GLEs. That analysis is a subject for another critical study. This section concerns the Grade 7, 2004 Math WASL. However, it is critical to understand that the GLEs are the driving forces of that assessment.

A tabulation of the Grade 7, Math GLEs yields a total of 185 specific learning objectives and that number alone is confounding to teachers as well. School is in session for 180 days. But, to be practical, there are really about 170 days of meaningful instruction. (Ask any teacher to verify that number.) If these GLEs are not phased in over grades 5, 6 & 7, then it means that every day one or more new math concepts must be taught in grade 7. Recall that the vast majority of these children are **NOT** at the formal level of cognition. There is absolutely no way students can master one new math concept per day. But the writers of the WASL "believe it to be." The operational term is "believe." With the OSPI having a school reform budget of $200,000,000 why were not experimental and control groups of schools set up to test the cognitive feasibility of those "beliefs?"

Analysis of the Grade 7 Math WASL

The Grade 7 2004 Math WASL is a paper and pencil test having 47 specific items. Table 8 presents a concept analysis of this assessment.

The algebra, probability, and geometric concepts being tested, for all reasonable placement, are 9^{th} or 10^{th} grade oriented. Please review Table 2 and Table 4. In the USA only 21% of the sampled 13-year-olds (8^{th} graders) could answer NAEP Level 300 items. And

TABLE 8. CONCEPT ANALYSIS OF GRADE 7 MATHEMATICS WASL 2004

Concepts	Item Numbers	Total Item in Category	Percent of Total
Algebra	3, 10, 13, (18), 27, 34, 38, 44, 46	9	14%
Arithmetic Skills	7, 12, 15, (16), 30, (39), (40), 45	8	12%
Estimation	2, 31, 32, 35, (42)	5	7%
Geometry	1, 6, (9), 16, 18, 23, 29, 41	8	12%
Graphing	5, 8, (9), (13), 14, 17, (20), (22), (27), (28), (32), 33, 36	13	20%
Multiple Part Responses	(5), 17, 19, 22, 24, 25, 26, (33), (36), (42)	7	11%
Probability	4, (5), 8, 14, 28, 37, 43	7	11%
Ratios	21, 47	2	3%
Weights and Measures	11, 20,	4	6%
Totals	**47**	**66**	**100%**

Table Notes:

1. Item numbers do not total 47 since items 5, 18, 39 and 40 appears to have multiple concepts, as do eight other items. All Graphing items should also be included in the multiple concept category.
2. The actual total number of items in the test is 47.
3. Twenty-three of the 47 items require extended responses (about 50%), thus are open to subjective scorer interpretation and a source of potential scorer error. Included are test items 5, 8, 13, 14a & b, 17a, b & c, 18, 20, 22a & b, 27, 28, 32, 33a & b, 36a & b, 38, 42a & b, 46.
4. Percent of total is computed on 66 items, thus there are overlapping values, due to several test items having multiple categories. The test items enclosed in parentheses ( ) overlap the conceptual categories.
5. All items require a high proficiency of reading. Recall the OSPI has reported that there is a correlation of 0.74 between the WASL reading and WASL math tests. The correlation could account for 50% of a student's variance! (One must ask if the math test is a reading test.)
6. Questions 1 and 6 both use the word "congruence." If a student misses item 1, then automatically the student misses item 6. This is an example of extremely poor test-item construction and test design.
7. A total of 11 of the 15 items requiring written responses relate directly to graphing.

only 1% could solve algebra items. With about one-fourth of 7th Grade Math WASL covering NAEP Level 350, one can easily understand why the failure rate is so high.

Conclusions

1. The 7th Grade Math WASL is in all reality a grade 9 test.
2. Test items do not progress from relatively easy to more difficult. They simply appear with no logical sequence.
3. A total of 9 math concepts are tested. Yet, 185 General Level Expectations are listed for Grade 7.
4. Reviewing the GLE's for grade 7 and 10 one observes parallel entries. That is, the grade 7 GLEs are almost identical in many cases to those of grade 10. Seventh graders are not "simple children." They are "simply children."
5. Regardless of the "paid for" praise (sole source contractors), the Math WASL is developmentally inappropriate for grade 7 children.

Policy Implications

There are critical instructional and policy implications associated with the findings of this analysis.

First, if the WASL tests are advanced beyond the mental cognition of grade 5 and 7 pupils, for most children, failure will be the ultimate end, regardless of what instructional techniques are used.

Second, what psychological impact will failing an inappropriate science and math WASL have on students and their ultimate attitudes towards science and math, and schooling in general?

Third, one may predict litigation by concerned parents and child advocacy groups against the State of Washington.

Fourth, scoring errors have been found nationally in virtually all mandatory high-stakes tests. These have led to class action law suits in which at least the state of Minnesota lost and paid out about $12 million to students and/or their parents due to scoring errors.

Fifth, the legislature is approaching fiscal irresponsibility or is not practicing fiscal accountability by continuing to fund the exorbitant WASL. (The current contract for the WASL with Pearson Educational Measurement is $70,800,000.) With the State of Washington predicting at least a $2.2 Billion state budget short fall, the massive $200,000,000 OSPI budget for school reform must be challenged.

Sixth, the legislature should commission an outside research organization to verify or refute this study. The legislature is specifically noted, not the Office of the State Superintendent of Public Instruction.

Seventh, any witness called to testify or to work on this topic must first be asked under oath, "How many sole source contracts or consultancies have you received from the OSPI, since 1993?" This simple question will eliminate paid for hire publicists--the "Armstrong Williams Effect."

APPENDIX A. A CRITICAL ISSUE NOT EXAMINED IN THIS REPORT

Appendix A illustrates the magnitude of an issue that is obviously beyond the scope of this report. Nevertheless, data presented in Table A.1 highlight a neglected topic with implications directly related to the *Washington Assessment of Student Learning*. The data in Table A.1, published by The Center on Education Policy in August 2003 and 2004, show the percentage of 10th grade students by ethnic group who succeeded passing the WASL *for first-time WASL test takers*. The legislature allows students to take the WASL **five times** until they pass or drop out of school. The issues associated with these WASL data must be considered a top priority to be addressed by the legislature, especially in light of "The No Child Left Behind Act of 2001" (PL107-110).

TABLE A.1. NINE ETHNIC GROUP PASS RATES ON THE 2002 AND 2003 MATHEMATICS AND READING WASL: GRADE 10 FIRST-TIME TEST TAKERS *(REPORTED IN PERCENTAGES OF THOSE MEETING THE ARBITRARILY SET STANDARD)*

Ethnic Group	***Math***		***Reading***	
	2002	**2003**	**2002**	**2003**
African Americans/Black	13	14	36	37
Alaskan Natives/Native Americans	21	22	44	43
Asian/Pacific Islanders	45	47	62	64
Latino/Hispanic	14	16	35	35
White/Caucasian	42	44	65	65
English Language Learners	9	8	13	12
Free or Reduced Price Lunch (Poverty)	19	24	39	43
Students With Disabilities	4	4	13	12
All Students	37	39	59	60

Source: *State High School Exit Tests Put to the Exam*. Washington, DC: Center on Education Policy, August 2003, page 133. (Let me add a footnote here. This study is worth its weight in time and effort. It may be downloaded at World Wide Web Site www.cep-dc.org.) **2003 Data Source:** *State High School Exit Exams: A Maturing Reform*. Center on Education Policy, Washington, DC: Table 3, p. 38, 2004.

The Author

Donald C. Orlich
PO Box 644237
Washington State University
Pullman, WA 99164-4237
Phone: (509) 335-4844 Email: dorlich@wsu.edu

Dr. Donald C. Orlich is professor emeritus at Washington State University. He has published over 100 professional papers and authored or co-authored over 30 monographs and books. He is the senior co-author of *Teaching Strategies: A Guide to Effective Teaching*, 7th Edition, Boston: Houghton Mifflin, 2004.

His specialty is curriculum and instruction, with expertise in science education as is evidenced with his funding of 22 National Science Foundation grants as Principal Investigator (PI) and Co-PI. Currently, he is Co-PI with Dr. Richard Zollars, Director of the School of Chemical Engineering at WSU on a $450,000 National Science Foundation teacher staff development project.

This paper is based upon an independent study done by the author, which was not sponsored by WSU. Washington State University encourages faculty to advance scholarship in their disciplines and strongly supports Academic Freedom.

Below is a selected list of the author's relevant publications.

________, (2000). "Education Reform and Limits to Student Achievement. "*Phi Delta Kappan,* Vol. 81, No. 6, pp. February, 468-472.

________, (2000). (Invited Paper). "A Critical Analysis of the Grade Four Washington Assessment of Student Learning." *Curriculum in Context*, Vol. 27, No. 2, Fall/Winter 2000, pp. 10-14. (This paper was awarded "Outstanding Affiliate Article Award" in Boston, March 2001 by the 170,000 member Association for Curriculum Development and Supervision.)

________, (2002). (Invited Paper). “Something Funny Happened on the Way to Reform,” *Northwest Professional Educator*, Vol. 1. No. 1, January 2002, pp. 1-2.

________, (2003, May 2). “A Nation at Risk—Revisited.” *Teachers College Record*. Retrieve at—http://www.tcrecord.orgID:11153.

________, (2003, June 12). “An Examination of the Longitudinal Effect of the Washington Assessment of Student Learning (WASL) on Student Achievement.” *Education Policy Analysis Archives*, Vol. 11, No. 18. Retrieve at--http://epaa.asu.edu/epaa/v11n18/.

________, (2004, Winter). “The Washington Assessment of Student Learning (WASL), Student Achievement and the No Child Left Behind Act.” *Leadership Information SIRS*, Vol. 3, No. 1, pp. 28-33. (School Information and Research Service, Olympia.)

________, (2004, September/October) (Invited Paper). “No Child Left Behind: An Illogical Accountability Model.” *The Clearing House*, Vol. 78, No. 1, pp.6-11.

________, (2005, In Press). With Glenn Gifford, “ The Relationship of Poverty to Test Scores.” *Leadership Information SIRS*, Olympia, Vol. 4. TBP.