BRT
behavioral research & teaching

PART 1 – VERTICAL SCALING PLAN
VERTICAL SCALING PROJECT – OREGON EXTENDED ASSESSMENT (NOVEMBER 24, 2014)
OPERATIONAL FIELD TEST APPROACH

Goal:

- **2015-16** - Establish a computer-based vertically scaled English language arts (combined reading and writing) and mathematics test. All tests will be linked to the Common Core State Standards (CCSS), Oregon Science Standards (ORSci), and/or Next Generation Science Standards (NGSS), as appropriate.
- Ensure the test has the possibility of being administered with paper-pencil format and that this administration is comparable to the computer administered format.

General Assumptions

- Test development will occur with (a) development of test item blueprint linked to CCSS in ELA and Math, to the ORSci and NGSS in Science; (b) completion of a technical specifications document; (c) development of an item writing training document; (d) completion of prototype items, and (e) collection of items from field-based teachers that are then standardized with graphics.
- Test linkage with CCSS/ORSci/NGSS is formally analyzed using another group of teachers to participate in a distributed item review study that collects information on alignment, bias (sensitivity), and perceived difficulty.
- Vertical scaling will only occur in grades 3-8 and only in English language arts and mathematics; Grade 11 and Science do not have contiguous grade levels (before or after) to support the development of a vertical scale.
- The 2014-2015 administration will be paper-pencil with scores entered into a secure BRT data entry website. The operational field tests will be fixed PDFs that teachers can access through BRT servers after successful completion of the online training and proficiency tests. This training will include a new module for training teachers on the access and administration of the new tests.
- Operational field-testing will occur throughout the test window (February 19, 2015 through April 23, 2015) and data will be used to calculate Annual Measurable Objectives (AMO).

- Each item appearing in operational field-test will have at least 200 student responses to be used in scaling the items both horizontally and vertically.
- A balanced design will be used, with grade level forms including both upper and lower grade level items (3rd grade level forms will only link up a grade level, while 8th grade forms will only link down a grade level).
- Each operational field-testing form will be composed of 48 total items, some of which are unique, some of which are used to anchor forms horizontally, and some of which are used to anchor the scale across grades.

BRT Computer Distribution in Field Testing

- BRT has developed a computer administration algorithm to ensure temporary (operational field-testing) forms have the maximum number of students taking the operational field test items. This algorithm seeds the 'form' available for the first *n* (e.g., 10) teachers to take the form and then forces the next *n* (e.g., 10) to take the next form. This pattern continues through all 5 forms and then begins again (beginning with 'form 1'). In this way, 20 waves of teachers take the test, ensuring that each 'form' has 200 students.
- Form distribution will be nested within teachers so that the same form is administered to all students for any given teacher before moving to another form. To provide comparability in count, teachers will need to sign up for an order, specifying the number of students to be given the assessment in each grade level and subject area.
- 200 students per form is expected, depending upon grade level frequencies; a secure file sharing system automatically assigns student the assessment form that ensures that form frequencies are balanced.
-

Because 200 students are needed for every form and grade level participation shrinks as grade level increases, there will be:

- 5 forms in grades 3-5,
- 4 forms in grades 6-8, and
- 3 forms in grade 11

BRT
behavioral research & teaching

3rd – 5th Grade Balanced Design in English Language Arts (combined reading and writing) and Mathematics
(5 forms)

Form												Total *New* Grade Level Items on Form		
1	$6A_5$	$6VS_{1A}$	$24U_1$	$6VS_{1B}$	$6A_1$									30
2					$6A_1$	$6VS_{2A}$	$24U_2$	$6VS_{2B}$	$6A_2$					30
3									$6A_2$	$6VS_{3A}$	$24U_3$	$6VS_{3B}$	$6A_3$	30
4	$6A_3$	$6VS_{4A}$	$24U_4$	$6VS_{4B}$	$6A_4$									30
5					$6A_4$	$6VS_{5A}$	$24U_5$	$6VS_{5B}$	$6A_5$					30
												Total *New* Grade Level Items = 150		

Note. U = unique items; V = vertically linked items (e.g., $6VS_1$ Grade 8 items from Form 1 will be included in Grade 7 unique items in Form 1, etc.); A = anchor items linked to next form. The 3rd grade forms will have 30U and only link up to 4th grade. Blue Shading = *New* items. Green Shading = *Verticallly Scaled* items (down one grade level) Yellow Shading = *Vertically Scaled* items(up one grade level)

5th Grade Science (not vertically scaled)
(5 forms)

Form												Total *New* Grade Level Items on Form		
1	$6A_5$		$36U_1$		$6A_1$									42
2					$6A_1$		$36U_2$		$6A_2$					42
3									$6A_2$		$36U_3$		$6A_3$	42
4	$6A_3$		$36U_4$		$6A_4$									42
5					$6A_4$		$36U_5$		$6A_5$					42
												Total *New* Grade Level Items = 210		

BRT
behavioral research & teaching

6-8th Grade Balanced Design in English Language Arts (combined reading and writing) and Mathematics
(4 forms)

Form												Total *New* Grade Level Items on Form		
1	$6A_4$	$6VS_{1A}$	$24U_1$	$6VS_{1B}$	$6A_1$									30
2					$6A_1$	$6VS_{2A}$	$24U_2$	$6VS_{2B}$	$6A_2$					30
3									$6A_2$	$6VS_{3A}$	$24U_3$	$6VS_{3B}$	$6A_3$	30
4	$6A_3$	$6VS_{4A}$	$24U_4$	$6VS_{4B}$	$6A_4$									30
														Total *New* Grade Level Items = 120

Note: The 8th grade forms will have 30U and only link down to 7th grade.

8th Grade Science (not vertically scaled)
(4 forms)

Form									Total *New* Grade Level Items on Form		
1	$6A_4$	$36U_1$		$6A_1$							42
2				$6A_1$	$36U_2$		$6A_2$				42
3							$6A_2$	$36U_3$		$6A_3$	42
4	$6A_3$	$36U_4$		$6A_4$							42
											Total *New* Grade Level Items = 168

BRT
behavioral research & teaching

11th Grade in English Language Arts (combined reading and writing), Mathematics, and Science (not vertically scaled)
(3 forms)

Form											Total *New* Grade Level Items on Form			
1	$6A_3$	$36U_1$	$6A_1$											42
2			$6A_1$	$36U_2$	$6A_2$									42
3					$6A_2$	$36U_3$	$6A_3$							42
														Total *New* Grade Level Items = 126

Note. U = unique items; A = anchor items linked to next form. Blue Shading = New Grade Level items

BRT
behavioral research & teaching

PART 2 – ITEM SAMPLING PLAN

THE FOLLOWING TABLES DEMONSTRATE THE BALANCE OF REPRESENTATION PLANNED FOR EACH TEST FORM ACROSS ALL RELEVANT DOMAINS IN ENGLISH LANGUAGE ARTS, MATHEMATICS, AND SCIENCE.

ENGLISH LANGUAGE ARTS

Domain	Grade 3	Used F1	Used F2	Used F3	Used F4	Used F5
RF	2	2	2	2	2	2
RI	4	4	4	4	4	4
RL	4	4	4	4	3	4
WR	4	4	4	4	4	4
LA	2	2	2	2	2	2
Needed	16	16	16	16	16	16

Domain	Grade 4	Used F1	Used F2	Used F3	Used F4	Used F5
RF	2	2	2	2	2	2
RI	4	4	4	4	4	4
RL	4	4	4	4	4	4
WR	4	4	4	4	4	4
LA	2	2	2	2	2	2
Needed	16	16	16	16	16	16

Domain	Grade 5	Used F1	Used F2	Used F3	Used F4	Used F5
RF	2	2	2	2	2	2
RI	4	4	4	4	4	4
RL	4	4	4	4	4	4
WR	4	4	4	4	4	3
LA	2	2	2	2	2	2
Needed	16	16	16	16	16	16

Domain	Grade 6	Used F1	Used F2	Used F3	Used F4
RI	5	5	5	5	5
RL	5	5	5	5	5
WR	4	4	4	4	4
LA	2	2	2	2	3
Needed	16	16	16	16	16

Domain	Grade 7	Used F1	Used F2	Used F3	Used F4
RI	5	5	5	5	5
RL	5	5	5	5	5
WR	4	4	4	4	4
LA	2	2	2	2	2
Needed	16	16	16	16	16

Domain	Grade 8	Used F1	Used F2	Used F3	Used F4
RI	5	5	5	5	5
RL	5	5	5	5	5
WR	4	4	4	4	4
LA	2	2	2	2	2
Needed	16	16	16	16	16

Domain	Grade 11	Used F1	Used F2	Used F3
RI	5	5	5	5
RL	5	5	5	5
WR	4	4	4	4
LA	2	2	2	2
Needed	16	16	16	16

MATHEMATICS

Domain	Grade 3	Used F1	Used F2	Used F3	Used F4	Used F5
OAT	7	6	5	6	5	5
NBT	2	1	1	1	1	1
NOF	3	1	1	1	1	1
MED	8	5	5	4	5	6
GEO	2	1	2	2	2	1
Total	22	14	14	14	14	14

Domain	Grade 4	Used F1	Used F2	Used F3	Used F4	Used F5
OAT	4	2	2	2	2	2
NBT	6	3	3	3	3	3
NOF	8	4	4	4	4	4
MED	5	2	2	2	2	2
GEO	3	1	1	1	1	1
Total	26	12	12	12	12	12

BRT
behavioral research & teaching

Domain	Grade 5	Used F1	Used F2	Used F3	Used F4	Used F5
OAT	3	2	2	2	2	2
NBT	8	4	4	4	4	4
NOF	6	3	3	3	3	3
MED	4	2	2	2	2	2
GEO	2	1	1	1	1	1
Total	23	12	12	12	12	12

Domain	Grade 6	Used F1	Used F2	Used F3	Used F4
GEO	3	1	1	1	1
RPR	3	1	1	1	1
TNS	9	4	4	4	4
EXE	6	3	3	3	2
STP	5	3	3	3	2
Total	26	12	12	12	10

Domain	Grade 7	Used F1	Used F2	Used F3	Used F4
GEO	3	2	2	2	2
RPR	2	1	1	1	1
TNS	7	5	5	5	5
EXE	2	1	1	1	1
STP	6	3	3	3	3
Total	20	12	12	12	12

Domain	Grade 8	Used F1	Used F2	Used F3	Used F4
GEO	4	3	3	3	3
TNS	2	2	2	2	2
EXE	6	4	4	4	4
STP	3	2	2	2	2
FUN	4	3	3	3	3
Total	19	14	14	14	14

Domain	Grade 11	Used F1	Used F2	Used F3
NAQ	2	2	2	2
ALG	2	1	1	1
FUN	7	5	5	5
GCO	2	1	1	1
GRT	1	1	1	1
GPE	3	2	2	2
GMG	1	1	1	1
STP	5	3	3	3
Total	23	16	16	16

SCIENCE

Domain	Grade 5	F1	F2	F3	F4	F5
LFS	5	4	5	5	4	4
PHS	5	5	5	4	5	5
ESS	5	5	4	5	5	5
ETS	2	2	2	2	2	2
Total	17	16	16	16	16	16

Domain	Grade 8	F1	F2	F3	F4
LFS	7	7	6	7	6
PHS	4	4	4	4	4
ESS	5	4	5	4	5
ETS	1	1	1	1	1
Total	17	16	16	16	16

Domain	Grade 11	F1	F2	F3
LFS	6	6	6	5
PHS	6	6	5	6
ESS	5	4	5	5
Total	17	16	16	16

PART 3 – ITEM REMOVAL DECISION RULES

IN ADDITION TO REVIEWING ITEM CONTENT, THE FOLLOWING DECISION RULES WERE USED TO DETERMINE WHETHER OR NOT ITEMS SHOULD BE MAINTAINED AS PART OF THE OPERATIONAL ITEM POOL. SELECTED ITEMS WERE USED TO DEVELOP THE VERTICAL SCALE, AS WELL AS FOR ALL STANDARD SETTING ACTIVITIES.

Classical Test Theory Review

ANCHOR ITEM RULES

Anchor items were removed if they had:

1. *p*-value - .95 and higher on all forms
2. Point biserial < .15 on either form

UNIQUE ITEM RULES

Unique items were removed if they had:

1. *p*-value - .90 and higher
2. Point biserial < .15

Rasch Model Analysis Review

Items were also removed if:

1. The outfit mean square was over 1.5 (exception: horizontal anchorss only removed if they were above 2.0)
2. Then, outfit mean square from 0-.5 on the other end of the spectrum.
3. Verify that all domains on all forms can still be represented