

**Oregon Extended Assessment – Achievement Level Descriptors
Development Process**
Thursday, May 21, 2015

Definition

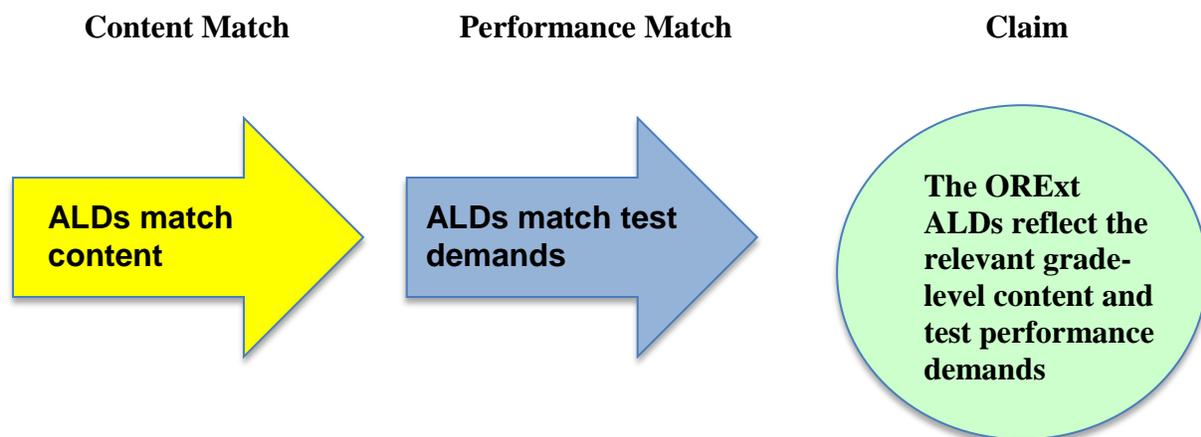
Achievement level descriptors (ALDs) "attempt to convey and more completely describe what performance within a particular category connotes about the test taker so classified." (Cizek & Bunch, 2007, p. 46). ALDs provide short summaries of what performance at a particular level on a given assessment means by describing what students should know and be able to do based upon test performance. When incorporated into the test development process appropriately, ALDs support a construct versus task validity approach to test design (Messick, 1995).

Intended Uses

ALDs have three intended uses: (a) test development, (b) standard setting guidance, (c) score interpretation (Egan, Schneider, & Ferrara, 2012). Ideally, ALDs are developed early in the test development process and guide all test design and development considerations, though using ALDs for test development has been less common in practice (Perie, 2008). ALDs are utilized to assist standard setting by contextualizing the quantitative significance of selected cut scores (Plake, Huff, & Reshetar, 2010). They are used thereafter on student score reports to help stakeholders understand what the student's test performance means by providing interpretive value for users. "Well-written PLDs [ALDs] disseminated in a timely manner can impact not only decisions about test development and cut scores, but also can inform teachers, parents, and students of the knowledge and skills intended to be learned in a year" (Perie, 2008, p. 24). To meet all three intended uses, ALDs must begin with the development of the test, then be used to guide standard setters, and finally help stakeholders interpret test results.

Claim to be Supported with ALDs

For the ALDs to provide a direct evidentiary statement for claims of knowledge and skill sets, they must be aligned to both test performance demands and content standards (Hambleton, 2001). The main claim is that the ORExt ALDs appropriately reflect the content of grade level standards in grades 3-8 & 11 in English language arts (ELA) and mathematics (Math) and in grades 5, 8, & 11 in science. In addition, the ALDs are reflective of what the assessment is designed to measure at each of those grade levels in four performance categories: (a) Level 1, (b) Level 2, (c) Level 3, and (d) Level 4.



Content Match

The Essentialized Assessment Frameworks (EAFs) are linked to the Common Core State Standards (CCSS) in English language arts and mathematics and in science to both current Oregon Science Standards (ORSci) and Next Generation Science Standards (NGSS). The essentialized standards that make up the EAFs, developed in the summer of 2014 by assessment experts at Behavioral Research & Teaching, are based on priorities of CCSS, ORSci, and NGSS grade-level standards. EAFs were subsequently reviewed by Oregon Department of Education staff. Oregon educators verified target standard priorities and degree of linkage with all standards. Educators rated the linkage of each essentialized standard to the target CCSS/ORSci/NGSS (2 = strong link, 1 = sufficient link, 0 = no link), as well as validated the

selection of the ES (yes/no). Complete results of the linkage study will be published in the 2015 technical report. In the end, the ALDs for the ORExt are direct distillations of the knowledge and skill statements found in the EAFs, and therefore distillations of the grade level content standards that have been reduced in terms of depth, breadth, and complexity.

Performance Match

All items on the 2015 ORExt included in the spring assessment have demonstrated sufficient alignment to the essentialized standards and therefore linkage to grade level content, with linkage ratings across the tested grade levels ranging from 1.73 to 1.82 in English language arts, 1.95 to 2.00 in mathematics, and 1.92 to 1.98 in science on the referenced 2-point scale. In addition, educators agreed with the essentialized standard selection process, with the percentage agreement in English language arts ranging from 94.7% to 96.7%, from 79.0% to 99.0% in mathematics, and 95.8% to 100% agreement in science. An example essentialized standard is provided in Table 1.

Table 1.

Example Essentialized Standard – Math Grade 3 Operations and Algebraic Thinking

Grade level content standard	Essentialized standard	Complexity level descriptors
(M.3.OA.1.3) Use multiplication and division within 100 to solve word problems in situations involving equal groups, arrays, and measurement quantities, e.g., by using drawings and equations with a symbol for the unknown number to represent the problem.	Solve word problems involving addition involving 1-20 and multiplication using 1- 5	<p>L- add 1-10</p> <p>M - add 11-20; multiply 1-2 by 2-4</p> <p>H - multiply 3-5 by 3-5</p>

ALD Development Process

Level 4 ALDs, the highest achievement level, were derived from high difficulty descriptors. Level 3 ALDs were derived from the medium difficulty descriptors. Level 2 ALDs were derived from the low difficulty descriptors. Level 1 ALDs were statements that the student demonstrated little to no performance of the knowledge and skills outlined in Levels 2-4. Where content themes emerged, the ALDs were collapsed in order to avoid redundancy. The ALDs can aptly be considered as abbreviated summaries of the tested content.

It is advised that all ALDs be drafted prior to any standard setting meeting (Cizek & Bunch, 2007), which is what has been done for the ORExt. A timeline of the 2015 ALD development process is provided in Table 2.

Table 2

ORExt ALD Development Timeline

Date	Description
April 2014	Essentialization process developed to elaborate grade level standards that have been reduced in terms of depth, breadth, and complexity at three levels: a) low difficulty, b) medium difficulty, and c) high difficulty.
August 2014	EAFs developed in ELA, Math, and Science in grades 3-8 & 11 at three levels of difficulty; designed for level H item content maps to Level 4, M item content maps to Level 3, L item content maps to Level 2, and students who are unable to meet item demands perform at Level 1
September – October 2014	ORExt items developed that align to the essentialized standards in the EAF documents
November 2014	ORExt items reviewed for alignment and bias; EAFs reviewed for linkage to grade level content standards (2 = Strong link, 1 = Sufficient link, 0 = No link)
December 2014 – February 2015	ORExt test forms developed based on test blueprints and complexity balance
April 2015	DRAFT ALDs developed in ELA, Math, and Science by ODE and BRT for review by Oregon State Board of Education; ALDs are distillations of EAF essentialized standards (L/M/H), aligned to content assessed
May 21, 2015	Presentation of DRAFT ALDs to Oregon State Board of Education
June 15-17, 2015	Standard Setting for the ORExt
June 25, 2015	Presentation of final cut scores to Oregon State Board of Education
July 10, 2015	Publication of final cut scores, ALDs, and Annual Measurable Objective reports for the ORExt

References

- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA.: Sage Publications.
- Egan, K. L., Schneider, C., & Ferrara, S. (2012). Performance level descriptors: History, practice, and a proposed framework. In G. J. Cizek (Ed.), *Setting Performance Standards* (pp. 79-106). New York: Routledge.
- Hambleton, R. (2001). Setting performance standards on educational assessments and criteria for evaluating the process. In G. J. Cizek (Ed.), *Setting Performance Standards: Theory and Applications* (pp. 89-116). New York: Routledge.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749.
- Perie, M. (2008). A guide to understanding and developing performance level descriptors. *Educational Measurement: Issues and Practice*, Winter 2008, 15-29.
- Plake, B. S., Huff, K., & Reshetar, R. (2010). Evidence-centered assessment design as a foundation for achievement level descriptor development for standard setting. *Applied measurement in education*, 23, 342-357. doi: 10.1080/08957347.2010.510964