ISAT Alignment with Idaho Content Standards

The Question: Are the ISAT assessment tests aligned with Idaho Content Standards?

The Importance of the Matter: The "No Child Left Behind" Act of 2001 as well as standard psychometric practice demands that an assessment test measure that which it is intended to measure.

Good alignment simply means that all of the test items match the state Content Standards well. A good match means that the test assesses the depth and breadth of the Content Standards. In other words, the questions have a range of difficulty levels and the test items measure a good range of specific knowledge and skills within the standards.

Methodology of Research: Some information about alignment was obtained from NWEA and has been reported in the discussion about content validity. Because of the importance of the matter, HumRRO was required to conduct an independent analysis of test alignment. HumRRO utilized the Webb alignment methodology and retained the services of Idaho educators to conduct the analyses. Other alignment methods exist, and they may have produced somewhat different results. However, these methods assume that the test is derived from the blueprint and the blueprint is derived from the standards. Since there is an issue with this relationship between the ISAT and the standards, it is highly likely that these other methods would have similar findings.

The alignment process used to analyze the ISAT was based on four measurements: categorical concurrence, depth of knowledge, range of knowledge, and balance of representation.

Categorical Concurrence is a basic measure of alignment between the Content Standards and test questions. Essentially, the percent reported in this category illustrates how many of the Content Standards were assessed by at least six test questions each.

Depth of Knowledge (DOK) measures the amount of cognitive complexity required by the test questions and by the standards separately. For example, if a part of the standard specifies that student should be able to simply recall basic knowledge, then the corresponding test item should be written to assess basic recall only. There are four levels of complexity ranging from recall to complex reasoning and thinking. Each specific content objective (or substandard) within a standard is assigned a DOK level and each test item is assigned a DOK level separately. Then, once the test item is matched to the particular standard, it is evaluated to ensure that the test item DOK level is either at or above the specific content objective. The methodology used (Webb, 1999) suggested that at least 50% of the items be at or above the matched knowledge, skill or ability's DOK.

Range of Knowledge can also be referred to as the "breadth" of how many of the objectives within the standards are covered by at least one test question. Another way of referring to range is how completely the test items cover the content in the Content Standards.

Balance of Representation takes the range of knowledge test one step further. This test analyzes the content objectives. For those objectives that are assessed, an approximately equal number of questions should be included in the test. The balance measure indicates whether the test questions really are distributed evenly across those objectives.

Synopsis of Findings:
As noted above, alignment using Webb's methodology is determined by four different measurements and therefore a single statement of "aligned" or "not aligned" is not possible. In other words, alignment is a matter of degree and does not provide a single yes or no answer. A test could be well aligned in some areas but not in other areas. The report highlighted several areas where the ISAT process can be improved and should be revised. For example, the reading blueprint has a category for vocabulary but there is not a specific Content Standard for vocabulary.

The study revealed that there are alignment issues that primarily stem from the lack of agreement between the Idaho Content Standards, the items obtained from NWEA, the "Reporting Goals" created by NWEA, and the test blueprints. These matters were reported in the section on "content validity."

The Idaho Math Content Standards have a large list of standards for this one subject area, which makes it difficult to fully assess the range of knowledge within those standards. The Language Arts/Communication Content Standards include five standards: reading, writing, speaking, viewing and listening. Thus, the ISAT for reading is built on just one standard, reading. The language usage test is built on just the writing standard. This small number of standards for a particular test means that it is more likely to have six items per standard (the categorical concurrence test), but that the balance of both the reading and the language usage test is not sufficient. If the reading and writing standards are viewed at the next level of granularity, it becomes evident that the tests are so focused on one or two of the sub-standards that the test does not do an adequate job of covering all the aspects in the standards.

In addition, the educators who participated in the study expressed concern about the wording of some of the Idaho Content Standards and that there were questions included in the test that were not directly related to a particular standard. For example, there were questions about the vocabulary word test items.

Implications for Future Direction: The alignment study confirmed weaknesses seen in the separate analysis of content validity. The problems stem from the use of the existing NWEA test structure, or blueprint, rather than the creation of an assessment made to specifically measure the Idaho Content Standards.

It is recommended that several actions be taken.

1. The Idaho Content Standards should be reviewed and possibly reorganized to solve the problems and inconsistencies that have been revealed. The review should seek balance across the content areas and creation of standards and skills that allow measurement of students across a wide range of proficiency.

2. The test blueprint should be reviewed and revised to reflect measurement of the reorganized Content Standards. If there is to be any content that is not routinely measured by the ISAT, this should be made explicit.

3. Test item specifications should be prepared to define how the Content Standards will be measured across the grade levels to address the lack of breadth in the ISAT.

4. Available items should be classified according to the content standard and grade level skill that they measure. Gaps in coverage should be noted and new items developed.

5. The use of "Reporting Goals" should cease and all reporting should be done in terms of the Idaho Content Standards.

6. The adaptive test structure and the use of "blended tests" should be reviewed in terms of the results of the HumRRO studies and the anticipated changes in items 1-5 above. Decisions must be made about the assessment structure and its capability to provide information needed by the State, by NCLB, and by classroom teachers.

Report: Idaho Standards Achievement Test (ISAT): Test Alignment Study, Task 3.1