

Just Completed!

The Consortium for Alternate Assessment Validity and Evaluation Studies (CAAVES Project 2006-2009)

The Consortium for Alternate Assessment Validity and Evaluation Studies (CAAVES project), an Enhanced Assessment Grant for the U.S. Department of Education's Office of Elementary and Secondary Education, was recently completed. This project was co-directed by Elizabeth Compton of the Idaho Department of Education and Stephen N. Elliott in Peabody College's Special Education Department and involved state directors of assessments from Arizona, Hawaii, Idaho, Indiana, Mississippi, and Nevada, the Discovery Education Assessment (DEA) team of Porter Palmer and William Dycus, and a research team of Ryan Kettler and Peter Beddow (Vanderbilt) Andrew Roach (Georgia State University), Michael Rodriguez (University of Minnesota), and Daniel Bolt (University of Wisconsin). These 13 project co-investigators collaborated to achieve three goals: (1) conduct a common validation study of the alternate assessments for students with severe cognitive disabilities in six of the states, (2) conduct an experimental study of item modifications on grade-level achievement tests for students with IEPs and persistent academic difficulties in four of the states, and (3) disseminate information to other state assessment leaders and special educators. The project successfully achieved all its goals. At its conclusion in March 2009, the CAAVES research team had worked with over 1,200 students with disabilities across six states and produced four research reports, presented eight papers at national meetings, and numerous deliverables produced by the principal investigators.

The CAAVES validation study of 1% alternate assessments was a six-state replication of a study published by [Elliott, Compton, and Roach \(2007\)](#). The intent of this study was to evaluate the alternate assessment for students with significant cognitive disabilities from each of the participating states, by using a multi-trait, multi-method framework with validated measures of adaptive behavior, academic competence, and academic achievement. Results from the study indicated that the constructs -- reading and mathematics - measured by the 1% assessments varied greatly from state to state, but in general tended to share about 50% of their variance with measures of adaptive behavior, as well as about 50% of their variance with measures of academic skills.To read more about this study [[link to the Kettler et al. report](#)]

The CAAVES experimental study of 2% or modified alternate assessments examined the reliability and validity of tests produced through modifications at the item level, and was inspired by a [recent policy](#) change allowing a small proportion of students with disabilities who persistently fail to meet proficiency on state achievement tests can do so through a test with modified items and achievement standards. For this study, items from DEA's online reading and mathematics tests for eighth graders were used and modified. The most common modifications included removal of a response option, simplification of language, addition of graphic support, and reorganization of layout. Before using the items with the entire sample, a cognitive lab study was conducted with a subsample of students to get their input into the utility and potential effectiveness of the item modifications (read Roach et al. report to get more information about this cognitive lab). The main study involved 755 students from four states and three eligibility categories, who completed items in Original, Modified, and Modified with Reading Support conditions. Results indicated the reliability of test forms was maintained throughout the modification process, that students who would be eligible for the assessment experienced a significantly greater gain in terms of item difficulty than did students who would not be eligible, and that shortening of an item stem was one particularly effective modification.

One unexpected deliverable from the CAAVES experimental study of 2% assessments was the *Test Accessibility and Modification Inventory (TAM)*, a set of guidelines developed by Special Education doctoral student Peter A. Beddow, along with LSI Coordinator of Data Services Ryan J. Kettler and project director Stephen Elliott. The inventory is a set of guidelines based on universal design, cognitive load theory, and sound item-writing principles, designed to measure and help improve the accessibility of items and tests for all students. This instrument was used by state assessment teams to guide the modification of the reading and math items. The *TAM* is a free tool that is perceived as having utility for educators, test-makers, and researchers interested in making their assessments as accessible as possible.

For information on an ongoing related project that extends research on item modification, visit <http://peabody.vanderbilt.edu/x8314.xml>