

*T A M I*

Test Accessibility and Modification Inventory™:

**Advancing a Technology of  
Universally-Designed  
Assessments**

**Peter A. Beddow, Ryan J. Kettler, & Stephen N. Elliott  
Vanderbilt University**

# TEST ACCESSIBILITY AND MODIFICATION INVENTORY: Advancing a Technology of Universally-Designed Assessments

The *Test Accessibility and Modification Inventory*<sup>TM</sup> (*TAMI*; Beddow, Kettler, & Elliott, 2008) is a decision-making tool designed to facilitate a comprehensive analysis of new and existing tests and test items with a focus on accessibility. Accessibility is defined as the extent to which an environment, system, or product eliminates barriers and permits equal access to all components and services for all individuals. The purpose of the *TAMI* is to assist test developers and educators with the design and modification of tests and test items to enhance their accessibility for all students. The validity of inferences that can be made from the scores of a test are directly related to its accessibility; thus, to the extent that the accessibility of a particular test is increased for the population for whom it was designed, the resulting scores will be more consistent, more precise, and more meaningful across the range of that population.

The *TAMI* was influenced by principles of universal design (e.g., the Center for Universal Design, 1997), test accessibility (e.g., Johnstone, Thurlow, Moore, & Altman, 2006), cognitive load theory (e.g., Clark, Nguyen, & Sweller, 2006), and fairness (Educational Testing Service, 2006). Research on testing accommodations, item writing, and item modification further influenced the content and organization of the instrument (e.g., Hollenbeck, 2002; Rodriguez, 2005). By applying the *TAMI* systematically, barriers can be reduced or eliminated and tests can be improved for many students for whom existing assessment instruments currently do not permit unfettered access.

## Organization and Use

The *TAMI* consists of 86 accessibility descriptors across two primary sections: Item Analysis (51 descriptors) and Computer-Based Test Analysis (35 descriptors) as listed in Table 1. The primary organization of the Item Analysis section is based on the five key elements of a test item: Passage / Item Stimulus, Item Stem, Visuals, Answer Choices, and Page & Layout. The Item Analysis section also contains a Fairness category, a set of considerations for analyzing items with respect to fairness to individuals and groups. The second section of the *TAMI* is divided into four categories based on key dimensions of computer-based assessments: Test Delivery System (i.e., login, user selection, test selection, navigation, response selection, etc.), Test Layout, Training, and Audio. Each category contains a set of descriptors to facilitate the development of accessible computer-based or online tests.

Table 1.  
*The Organization of the TAMI*

Item Analysis	Computer-Based Test Analysis
<b>1. Passage / Item Stimulus</b> (10 descriptors)	<b>1. Test Delivery System</b> (19 descriptors)
<b>2. Item Stem</b> (10 descriptors)	<b>2. Test Layout</b> (5 descriptors)
<b>3. Visuals</b> (11 descriptors)	<b>3. Training</b> (2 descriptors)
<b>4. Answer Choices</b> (6 descriptors)	<b>4. Audio</b> (9 descriptors)
<b>5. Page / Layout</b> (9 descriptors)	
<b>6. Fairness</b> (5 descriptors)	

The *TAMI* utilizes Accessibility Worksheets that correspond to each section of the inventory. These worksheets provide a team of users a consistent framework for documenting their analyses of items and recommended revisions. To analyze items, raters use the Item Accessibility Worksheets to record an accessibility rating for each category, using a 4-point scale (0 = not accessible; 1 = minimally accessible; 2 = moderately accessible; 3 = maximally accessible). The ratings are then summed to yield a Total Accessibility Score which can be used to facilitate comparison across a large item pool. Additionally, the worksheet can be used to record recommended item modifications and to document final changes to items. Similarly, to analyze computer-based tests, raters use the Computer-Based Test Accessibility Worksheets to record their categorical ratings from the corresponding section of the *TAMI*, recommend modifications, and document final changes.

## Field Test

In July of 2008, the authors of the *TAMI* conducted an examination of the instrument in a large western state in the United States. The *TAMI* was presented as part of training exercise given by a team of researchers from Vanderbilt University and the University of Minnesota to prepare a group of educators, content-area experts, and assessment specialists to modify a pool of items for use in a large-scale assessment for students who have failed to make adequate progress given sufficient grade-level instruction and who have had a consistent record of below-proficient performance on the state's regular grade-level assessment. Recent regulations under the *No Child Left Behind Act* have permitted states to develop these new assessments, called alternate assessments based on modified achievement standards (AA-MASs), and to report up to 2% of students in a state proficient in reading and mathematics using the tests.

Each participant ( $N = 32$ ) was randomly assigned to analyze reading and math items in one of two activity packets using the *TAMI*. Half of the items in each of the packets were unmodified items out of a pool of formative assessment items from Discovery Educational Assessment (DEA). The other half of the packet contained modified versions of the original items. During the following school year, the effects of the modifications were examined in an experimental study with over 700 students across the four states. Data provided valuable insights about effective modification procedures and contributed to the development of the *TAMI* (for more detail on this study, see the following report: Kettler, Rodriguez, Bolt, Elliott, Beddow, Kurz, et al., 2008 or visit <http://peabody.vanderbilt.edu/tami.xml>).

The two packets were balanced with respect to original and modified items; each group received modified versions of the items received by the other group in original form, and vice versa. This design permitted the comparison of *TAMI* Total Item Accessibility Scores between original and modified versions of the same items to evaluate the sensitivity of the instrument. Results of this initial validity study indicate the *TAMI* is able to be used by educators to detect differences between items that have been modified with a focus on accessibility and those that have not, even when raters have received minimal training.

Additionally, all participants and state assessment leaders ( $N = 36$ ) completed feedback forms on which they were asked to rate the instrument on a 3-point scale (1 = low/poor; 2 = moderate/acceptable; 3 = high/excellent). Results indicated highly favorable perceptions of the *TAMI*'s usefulness for item modification work ( $M = 2.70$ ), organization ( $M = 2.78$ ), completeness of category descriptors ( $M = 2.88$ ), utility for documenting modifications (mean = 2.71), and overall usefulness for the current work ( $M = 2.78$ ). A number of participants questioned the feasibility of assigning ratings to items for a large item pool ( $M = 2.40$ ), but several commented that the process was useful for item modification training. Most indicated that they would be willing to use the *TAMI* again ( $M = 2.72$ ) and several participants noted that they planned to use the instrument in their classrooms to assist in designing curriculum-based assessments and for future item modification work.

## Conclusion

Analyzing tests and test items with a focus on enhancing their accessibility is essential for developing assessments that are intended for use with students with a broad range of abilities and needs. Grounded in research and theory, the *TAMI* was expressly designed to facilitate the process of developing and modifying tests and test items that are accessible to as many students as possible. By using the instrument to systematically increase access to new and existing tests, the practice of testing can be improved.

## References

- Beddow, P. A., Kettler, R. J., & Elliott, S. N. (2008). *Test Accessibility and Modification Inventory (TAMI)*. Nashville, TN: Vanderbilt University. <http://peabody.vanderbilt.edu/x8312.xml>
- Center for Universal Design. (1997). *The principles of universal design*. Retrieved August 4, 2008 from <http://www.design.ncsu.edu/cud>
- Clark, R., Nguyen, F., & Sweller, J. (2006). *Efficiency in learning: Evidence-based guidelines to manage cognitive load*. San Francisco, CA: Pfeiffer.
- Johnstone, C., Thurlow, M., Moore, M., & Altman, J. (2006). *Using systematic item selection methods to improve universal design of assessments* (Policy Directions 18). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Kettler, R. J., Rodriguez, M. C., Bolt, D. M., Elliott, S. N., Beddow, P. A., Kurz, A., et al. (2008). Modified multiple-choice items for alternate assessments: reliability, difficulty, and the interaction paradigm. Unpublished manuscript. Peabody College of Vanderbilt University.
- Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice*, 24, 3-13.

The *TAMI* was developed by Peter A. Beddow, Ryan J. Kettler, and Stephen N. Elliott of Vanderbilt University as part of a USDE Enhanced Assessment Grant (S368A060012) project in the Consortium for Alternate Assessment Validity and Experimental Studies (CAAVES). Resources on universal design and test fairness from the National Center on Educational Outcomes and the Educational Testing Service were influential in the development of this item design and evaluation tool.

For more information about the development of the *TAMI* and to request copies, please visit the *TAMI* web page at the Learning Sciences Institute, Vanderbilt University:

<http://peabody.vanderbilt.edu/tami.xml>

