

The Development and Use of an Item Accessibility and Modification Guide

Peter A. Beddow, Ryan J. Kettler, & Stephen N. Elliott
Peabody College of Vanderbilt University

Constructing items for modified achievement tests that ensure the test achieves its stated purpose is a challenging task. In addition to yielding scores that meet the requirements of acceptable validity and reliability, Title I regulations require that assessments include maximally readable, comprehensible, and legible items (34 C.F.R. §200.2(b)(2)), and use principles of universal design (34 C.F.R. §300.160(f)). Several professionals have suggested using modified items from the general assessment, reducing item complexity, and decreasing the number of response choices from four to three are reasonable strategies for developing an alternate achievement test for modified content standards (Burling, 2007). To respond adequately to these and other concerns, the generation and modification of test items for state modified achievement tests demands intensive cooperation among professionals with expertise in five primary areas: (a) teaching students with disabilities, (b) grade-level academic content, (c) universal design principles, (d) research-based item writing tactics, and (d) information processing and responding skills of students with disabilities.

In preparation for the development of a pool of items for a modified achievement test, the leadership of the Consortium for Alternate Assessment Validity and Experimental Studies (CAAVES) drafted the *Item Accessibility and Modification Guide (IAMG)* (Beddow, Kettler, & Elliott, 2007). The tool was designed “to facilitate the analysis of test item features with the purpose of enhancing meaningful responses from, and providing equal access for, all students” (*IAMG*, p.1). The *IAMG* was expressly influenced by principles of universal design and accessibility. In addition to consulting research on testing accommodations, item writing, and item modification, the authors used publications by the Center for Universal Design (1997) and documents from National Center for Educational Outcomes (Johnstone, Liu, Altman, & Thurlow, 2007; Johnstone, Thurlow, Moore, & Altman, 2006; Thompson, Johnstone, Anderson, & Miller, 2005) to shape the organizational and content-related features of the Guide.

The Center for Universal Design (CUD) has recommended the consideration of seven principles for the universal design of environments, products, and services “to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design”: (a) equitable use; (b) flexible use; (c) simplicity and intuitiveness; (d) perceptible information; (e) tolerance for error; (f) low physical effort; and (g) size and space for approach and use (CUD, 1997, p.1). All seven principles and their respective guidelines were considered in the development of the instrument; however, given the minimal physical demand and space requirements of most assessments of academic standards, the first five were primarily used to guide the inclusion of items on the *IAMG*. As a result, the *IAMG* helps modify items along the lines of Cognitive Load Theory (Clark, Nguyen, & Sweller, 2006), reducing the extrinsic cognitive load so that the student can better show what she or he knows.

The *IAMG* is divided into two main sections containing categorized checklist items: (a) Considerations for Universally Designed Assessment Items and (b) Computer-based Tests: Considerations for Universally Designed Items. Several category subscales are contained within each section, each of which is designed to focus raters on specific features of assessment items. Following completion of the checklist, the rater is prompted to provide a summative evaluation of the accessibility of the item for that category. At the

conclusion of each section, evaluations can be summed across categories to yield an overall Item Accessibility Rating for that section. Although this summation provides a simple way to categorize each item as having high, medium, or low accessibility, more research needs to be done on the additive nature of these considerations, before that total can be interpreted as a definitive indicator of accessibility (i.e., some considerations may be more or less important for certain types of tests and items). The *IAMG* is intended only for use as an assistive tool for evaluating universal design and accessibility features of items and tests. It is not a norm- or standard-referenced measure and was not designed to provide definitive analyses for use in technical reports. Rather it provides item developers a common lens to examine important features of items and tests, and thus facilitates the collaborative work of educators who wish to develop objective measures of modified achievement standards.

Conclusions

Based on feedback forms completed by the CAAVES participants, the initial reception of the *IAMG* was positive. The majority of feedback indicated that the Guide helped groups become aware of the range of possible modifications for items and was useful in the evaluation of the subsequent pool of items during the 3-day working session. A number of workshop participants commented that the checklist was thorough and encompassed the necessary aspects of the modification process. Several users also suggested that the Guide could have utility for classroom teachers in analyzing their classroom-based assessments, or as a training tool for educators who are inexperienced in test development. Overall, feedback indicated that the *IAMG* provided a useful framework that may have future utility for educators, researchers, and test developers for conceptualizing and evaluating the accessibility and universal design of tests and test items.

References

- Beddow, P. A., Kettler, R. J., & Elliott, S. N. (2007). *Item Accessibility and Modification Guide*. CAAVES project. Unpublished rating scale.
- Burling, K. (June, 2007). NCLB regulations for modified achievement standards (2%): A white paper from Pearson Educational Measurement. Retrieved December 29, 2007 from <http://www.pearsonedmeasurement.com/downloads/white/wp0702.pdf>
- Center for Universal Design (1997). North Carolina State University. Retrieved July 15, 2007 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., Liu, K., Altman, J., & Thurlow, M. (2007). *Student think aloud reflections on comprehensible and readable assessment items: Perspectives on what does and does not make an item readable*. (Technical Report 48). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- 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.
- Thompson, S. J., Johnstone, C. J., Anderson, M. E., & Miller, N. A. (2005). *Considerations for the development and review of universally designed assessments* (Technical Report 42). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.