

Name: Stephanie W. Cawthon

Affiliation: Assistant Professor
School Psychology Program
Department of Educational Psychology

Address: The University of Texas at Austin
1 University Station, MC D5800
Austin, TX 78712

Email: Stephanie.cawthon@mail.utexas.edu

Phone: 512-471-0287

Fax: 512-475-7641

Cawthon Personal Research Agenda

The No Child Left Behind Act of 2001 (NCLB) raised the level of accountability for student performance by placing increased emphasis on results from statewide, standardized assessments. The high-stakes nature of this process makes it imperative that states develop credible, high quality assessment systems (Linn, 2000). However, even a quality assessment can be challenging for students with disabilities to access test content and demonstrate their knowledge and skill (Phillips, 1994). *Assessment accommodations* are often used in an effort to “even the playing field” and reduce barriers that are not related to test content (Lazarus, Thurlow, Lail, Eisenbraum, & Kato, 2006). However, there are some concerns that arise using accommodations because they may alter the *validity* of the assessment (Koretz & Barton, 2003). Reducing potential threats to validity is thus an important issue when making decisions about accommodations use in standardized assessments.

Previous Research Efforts

My previous research efforts focus on two main lines of investigation: (a) accommodations use with students who are deaf or hard of hearing (SDHH) and (b) the development of a Validity Evaluation Tool. The case of SDHH is interesting because these students come to the assessment process with a wide range of linguistic, literacy, and content area skills that are affected both by their disability and by the varying modes of communication used in their classroom environment, such as American Sign Language. Through national surveys and policy analysis, the SDHH assessment research has contributed to an ongoing dialog accommodations affect student performance on standardized tests. My second line of research is the development of the Validity Evaluation Tool (VET). This project grew out of the SDHH research findings that teachers want to consider multiple student and test characteristics when

choosing accommodations as well as research that demonstrates the differential effects of these factors. The goal of the VET tool is to systematically measure the impact of multiple factors on test validity such as the individual student's skill level in the test content area, linguistic complexity of the test item, and the format of the test. The intent of this tool is to facilitate informed decision making about accommodations as well as the policies that guide their use. Thus far, the VET team has conducted a literature review of empirical studies in the area, developed a draft VET measure, and is conducting a pilot study of VET components.

Proposed Research Questions

My personal research agenda for the next 3-5 years is to investigate two broad research questions: (1) What linguistic, academic proficiency, test content, and test item characteristics affect the validity of accommodations used by students with disabilities? (2) How might research findings about accommodations validity, using a measure such as the VET, be integrated into the accommodations decision-making process? The first step will be to address these questions for students with a Learning Disability (LD), a high-incidence population. The proposed multi-site collaborative study addresses this goal for elementary and middle school students.

Implications for Schools

Schools will benefit from this project in two ways. The first is in the opportunity to participate in the development of a tool designed to facilitate a challenging task: choosing valid accommodations. As partners in this process, schools will provide essential input into refining VET components. Second, the research findings and resulting VET tool will provide teachers a way to consider the multiple factors that affect how a student, test, and accommodation interact during testing. Increasing the validity of test scores will help students more meaningfully participate in statewide, standardized assessments, including those used for accountability.

References

- Koretz, D. & Barton, K. (2003). *Assessing students with disabilities: Issues and evidence*. Los Angeles, CA: Center for the Study of Evaluation.
- Lazarus, S. S., Thurlow, M. L., Lail, K. E., Eisenbraun, K. D., & Kato, K. (2006). *2005 state policies on assessment participation and accommodations for students with disabilities*. (Synthesis Report No. 64). Minneapolis: University of Minnesota, National Center on Educational Outcomes.
- Linn, R. L. (2000). Assessment and accountability. *Educational Researcher*, 29(2), 4-16.
- No Child Left Behind Act of 2001, 20 U.S.C. § 6301 (2002).
- Phillips, S. E. (1994). High-stakes testing accommodations: Validity versus disabled rights. *Applied Measurement in Education*, 7(2), 93-120.

Cawthon Collaborative Multi-Site Research Program Proposal

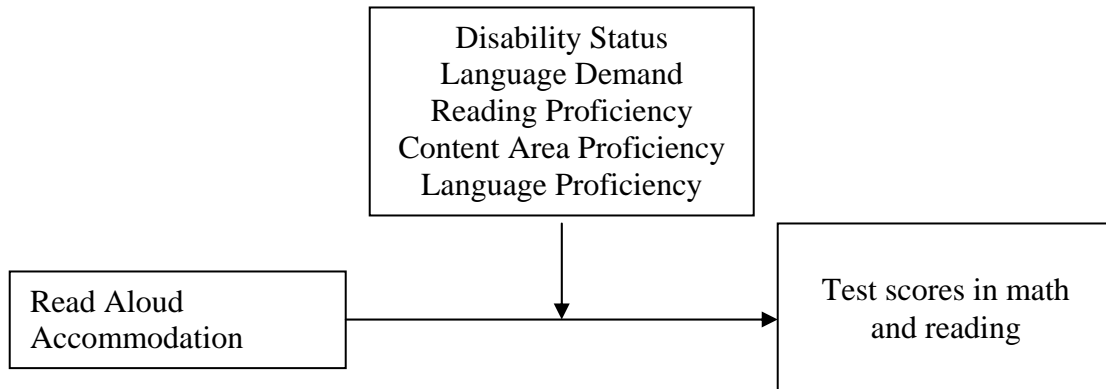
Issue: Students with disabilities often need assessment accommodations to participate in large-scale, standardized tests such as those used to evaluate academic progress under No Child Left Behind. Current state assessment policies provide only very general guidelines on how to ensure valid test scores. This is particularly true for accommodations that might change the difficulty of the assessment, such as having test items read aloud to students instead of participants reading items on their own (Read Aloud). The research literature on test score validity indicates that the benefits of accommodations may depend on specific factors such as the test subject, item complexity, or student characteristics. Yet few research studies incorporate the *multiple* constructs into a coherent measure of accommodations and their effect on test score validity.

Research Questions: The purpose of this study is to investigate the impact of multiple student and test characteristics on the validity of accommodated test scores. This study will use a “differential boost” definition of validity. An accommodated score will be viewed as valid if students with disabilities receive a greater increase in scores, or boost, than students without disabilities (SWOD). We will begin with students with a learning disability (LD) because they are the most prevalent group of students with disabilities enrolled in public schools. We will focus on the Read Aloud accommodation because it is widely used and yet it use faces consequences in many state assessment policies.

Research Question 1: Does the Read Aloud accommodation change math and reading assessment outcomes for LD students versus SWOD?

Research Question 2: Do student disability status, reading proficiency, content area proficiency, language proficiency, and language demand of the test items moderate the effect of accommodations on math and reading scores (Figure 1)?

Figure 1. Do Multiple Factors Affect Validity of Accommodated Test Scores?



Participants: Students from multiple sites will participate in this study. Participants will complete (a) a set of student characteristic measures (b) an abridged set of test items with and without the Read Aloud accommodation. Scores from the LD students will be compared with SWOD peers. To have the power to measure at least medium effect sizes, an effort will be made to recruit approximately 200 5th and 6th grade students, 100 LD and 100 SWOD, for the first year of the project (2009-2010). The project then expands its base in 2010-2011 by including LD and SWOD students in elementary grades (3rd and 4th) and middle school (7th and 8th). This breadth is necessary because the variables included in this model will likely vary by age of the students. The project therefore hopes to reach a total of 300 LD and 300 SWOD over two years.

Budget: Budget items for this project include salary for a 10-hour per week graduate research assistant, test materials, teacher compensation, office supplies and travel to school sites (see table 1 on following page). I have also applied for funds from the Richardson Policy Fellowship program that would contribute additional resources for the first year of the project (2009-2010). If the Richardson funds do not come through, the SPRCC funds will be used for the first year activities while money for the second year are sought from the U.S. Department of Education and foundations that support projects for students with learning disabilities.

Table 1. *Proposed SPRCC Project Budget*

Category	Description	Outside Funds (pending)	SPRCC Funds
<i>Research Assistant (salary, tuition, and fringe)</i>	10 hours per week for 2009-2010 and 2010-2011	2009-2010: \$9,000	2010-2011: \$9,000
<i>Research Expenses</i>	\$15 per student for test materials	200 sets: \$3,000	400 sets: \$6,000
<i>Program Activities</i>	\$25 for participating teachers	20 teachers: \$500	40 teachers: \$1,000
<i>Office Supplies</i>	Copying, filing folders, postage for measures, telephone calls, printing, etc.	-	\$1,000
<i>Travel to school sites</i>	Gas and related expenses	\$450	\$1,000
<i>Total Project Budget</i>		<i>\$12,950</i>	<i>\$18,000</i>

Advantages and Disadvantages: There are several advantages to conducting a multi-site study for this project. The first advantage is the opportunity to establish relationships with multiple schools and programs, including those that are designed to serve specific student populations that are located in various cities across the country. The second advantage is that it helps us meet the challenge of conducting research on students within a targeted age range by expanding the reach beyond one metropolitan area. Finally, collaboration allows us to draw on partners' expertise in study design, implementation, and analysis, strengthening the research process.

There are, however, a number of disadvantages to a multi-site study. It can be difficult to maintain clear and consistent communication between project PIs, grad students, and school site participants. This may be addressed, in part, with a clear research protocol that can be used reliably by multiple users. A second disadvantage is logistics: it is challenging to get materials to multiple school sites when they are dispersed across different settings. Finally, coordination and timing of research has a greater overhead of time and resources in a multi-site project. Strong organizational infrastructure and sufficient staff are thus critical to the success of this project.

Abstract

Students with disabilities often need assessment accommodations to participate in large-scale, standardized tests such as those used by states to evaluate academic progress. The research literature on test score validity indicates that the benefits of accommodations may depend on specific factors such as the test subject, test item linguistic complexity, or student characteristics. The purpose of this study is to investigate the impact of *multiple* student and test characteristics on the validity of accommodated test scores. Using both a differential boost and multiple regression analysis of results, this project will investigate whether student disability status, reading proficiency, content area proficiency, language proficiency, and/or language demand of the test items moderate the effect of “read aloud” accommodations on test scores. In its first year, the project focuses on late elementary grade (5th and 6th grade) students with a Learning Disability (LD). The project then expands its base in the second year by including students in elementary grades (3rd and 4th) and middle school (7th and 8th). This expansion is necessary because the student characteristics included in this model will likely vary by grade range. Students will be recruited from schools and programs across the country to meet the sample size needs for this study. The results of this project will be used to build the Validity Evaluation Tool, or the VET. The VET will provide teachers with research-based evidence for what accommodations may be valid for use, while taking into account the various student and test characteristics involved in the assessment process. Results from the proposed project will determine what recommendations the VET tool will make for the valid use of Read Aloud for LD students in math and reading standardized assessments.