2005

Extended Time Test Accommodation: Directions for Future Research and Practice

LoriAnn S. Stretch
Jason Osborne

Follow this and additional works at: https://scholarworks.umass.edu/pare

Recommended Citation
Stretch, LoriAnn S. and Osborne, Jason (2005) "Extended Time Test Accommodation: Directions for Future Research and Practice," Practical Assessment, Research, and Evaluation: Vol. 10 , Article 8. DOI: https://doi.org/10.7275/cs6a-4s02
Available at: https://scholarworks.umass.edu/pare/vol10/iss1/8

This Article is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Practical Assessment, Research, and Evaluation by an authorized editor of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
Extended Time Test Accommodation: Directions for Future Research and Practice

LoriAnn S. Stretch and Jason W. Osborne
North Carolina State University

Several pieces of legislation, most recently No Child Left Behind, hinge on effective assessment of students with disabilities. Mandated inclusion and mandated accountability for progress raises many interesting questions regarding how to fairly, effectively, and validly perform assessments on students with disabilities. The purpose of this article is to summarize and discuss current research on extended time testing, particularly the implications of ETTA for assessment. Research indicates that in regard to students with learning disabilities practitioners and researchers need (a) to find ways to provide realistic, not inflated, estimates of a student's ability, (b) to determine if test instruments are suitable for use with students with learning disabilities, (c) to consider whether students with learning disabilities were included in the normative and validation samples, (d) to know that scores from accommodated assessments should be tentative, and (e) to weigh whether scores from assessments that are not validated are more useful than information available from other sources.

Since the passage of the Americans with Disabilities Act of 1990 (ADA), the number of students with disabilities entering college has been increasing (Dalke & Schmitt, 1997; Fichten, 1988; Levinson & Ohler, 1998; Thomas, 2000; Vogel, 1987). More than 892,000 students with disabilities attended college in the U.S. in the 1995-96 academic year, and the number continues to increase (U.S. Department of Education, 2000). In spite of demographic trends confirming the efficacy of laws regarding access, students with disabilities have been less than successful in participating fully in the college experience and in attaining a college degree (Hall & Belch, 2000). The challenge for most colleges and universities is how to provide equal access to students with disabilities while maintaining university standards. Consequently, colleges and universities must strive to understand the specific factors that contribute to the academic success of college students with disabilities in order to provide access while maintaining university standards.

At the same time postsecondary institutions are wrestling with these issues, the passage of No Child Left Behind (NCLB) has forced public education to confront similar issues relating to assessment of students with disabilities, particularly how to validly combine assessment and accommodation. Therefore, the purpose of this article is to summarize and discuss the implications of extended time testing (one of the more common accommodations) for valid assessment of students and to propose future research and practice regarding extended timed testing.

**What is an Accommodation?**

An accommodation is generally a change in assessment materials or procedures that address
aspects of students’ disabilities that may interfere with the valid assessment of their knowledge and skills (Thurlow & Bolt, 2001). While both accommodations and modifications are nonstandard conditions, accommodations differ from modification in that modifications change what construct is being measured; accommodations should only change how a construct is being measured, with the goal of more valid assessment of students with disabilities, a "leveling of the playing field," allowing a student with a disability to demonstrate his or her "true" abilities to the same extent as students without disabilities (Elliott, Kratochwill, & McKevitt, 2001; Fuchs, 2001; Fuchs, Fuchs, Eaton, Hamlett, Binkley, & Crouch, 2000; Shriner, 2000; see also Elliott & Braden, 2000; McDonnell, McLaughlin, & Morison, 1997; Willingham et al., 1988). One challenge, however, is to accommodate without introducing new measurement biases or measuring a different construct completely.

**What are the Legal Issues Surrounding Accommodations?**

The assessment of students with disabilities has taken on considerable importance with the passage of the ADA which mandated inclusion of students with disabilities, and more recently NCLB which holds schools accountable for showing progress for all students, including students with disabilities. The purpose of Section 504 and ADA as applicable to educational institutions and testing programs was to ensure equal and effective access to education for qualified students with disabilities (Heaney & Pullin, 1998). Similarly, the ADA has the expectations of students and their parents that educational institutions and testing programs will provide the necessary accommodations to assist students in obtaining access to educational programs and completing their degrees.

Most admissions testing programs for post-secondary programs provide accommodations in their testing formats for students with disabilities in response to the enactments of Section 504 and ADA. Section 504 mandates that admissions tests administered to students with disabilities must be validated and that scores resulting from such instruments must reflect ability and aptitude rather than any disabilities extraneous to what is assessed (Geisinger, 1994). The ADA (1990) requires that “when an examination is administered to an individual with a disability…the examination results accurately reflect the individual’s aptitude or achievement level…rather than reflecting the individual’s impaired…skills” (36.309b(i)). Therefore, current social policy and the provisions of Section 504 and the ADA require that students with disabilities be afforded reasonable accommodations so that they might participate fairly in admissions testing (Geisinger, 1994; Geisinger & Carlson, 1995; Department of Justice [DOJ], 1996).

However, the word, “reasonable,” is ambiguous and differs depending upon the circumstances of the assessment and the institution administering the assessment. If a requested accommodation appears to violate the integrity of an assessment and interferes with the test’s intended purpose, then the accommodation is not legally required and may prove counterproductive (Zuriff, 1996-97). A prime example of the complexity of this situation is with timed tests. For some students with disabilities, tests with strict time limits may violate provisions of the ADA in that the timed tests may measure performance under conditions of the disability and consequently not allow students with disabilities to demonstrate their knowledge and skill (Zuriff, 2000). The ADA does not require time accommodations if speed of work is part of what is being evaluated by the assessment (Zuriff, 2000). However, in cases where time is not a critical aspect of the assessment, the students may qualify for extended time on the test. “If, as a function of a disabling condition, an examinee’s knowledge and skills cannot be fully demonstrated under standardized testing conditions, the obtained score will not accurately reflect the examinee’s level of achievement but, rather, the extent of the disability” (Munger & Lloyd, 1991, p.53).

**What Do Accommodations Do to Test Standardization?**

Every year, thousands of colleges and university applicants with learning disabilities present scores from standardized examinations as part of the admissions process for postsecondary education. Each year, millions of K-12 students in the U.S. take standardized tests, a significant percentage using some form of accommodation. With NCLB, practitioners must insure that those standardized scores are valid.
Unfortunately, the research on test accommodations is limited and fails to provide an adequate basis for definitive conclusions about how test accommodations affect students with learning disabilities (Fuchs, 2001) or how accommodations affect standardization (Anastasi, 1988; Willingham, 1991; Zurcher & Bryant, 2001). This is an increasingly important issue as the number of students identified as having learning disabilities continues to rise (Heaney & Pullin, 1998) and readers are encouraged to refer to Sireci, Li, and Scarpati (2003) for a thorough review of the literature.

**What Do Accommodations Do to Test Validity?**

Providing appropriate accommodations and ascertaining the consequences of accommodations present significant issues concerning the validity of the inferences made from scores on modified tests (Heaney & Pullin, 1998). Significant questions exist about fairness in testing students with disabilities and the validity of the inferences made from scores on modified versions of admissions tests (Fuchs, Fuchs, Eaton, Hamlett, Binkley, et al., 2000; Heaney & Pullin, 1998; Thurlow & Ysseldyke, 1995). As tests are accommodated, especially if the accommodations are locally determined (as opposed to being determined by test publishers), error may introduced into the testing process (Geisinger, 1994). There is disagreement about which accommodations preserve the meaningfulness of scores for students with disabilities (Fuchs, Fuchs, Eaton, Hamlett, Binkley, et al., 2000). Many accommodations currently used to address the disadvantages inherent in the learning disabled population (e.g., extended time, decoding questions, encoding responses) may actually distort the meaning and interpretation of scores. Scores from tests with nonstandard administrations for students with learning disabilities are more comparable to those of students without learning disabilities but are of questionable validity for their intended purpose, such as assisting college admissions officers to select from among applicants (Zurcher & Bryant, 2001). The Willingham et al. (1988) study demonstrated lower predictive validity of test scores from nonstandard administrations. The authors recommended that (as with all students) multiple data points are needed when making important decisions (e.g., admissions or selection decisions).

**Extended Time Testing Accommodation (ETTA)**

ETTA is one of the most common accommodations on standardized tests in the US. According to Ragosta & Wendler (1992, p. 100) students with disabilities taking the SAT under ETTA receive up to twice the standard examination time. However, limited research has been conducted to examine the effect of testing time on the performance of students with learning disabilities or to determine the amount of time actually needed by students with learning disabilities (Munger & Lloyd, 1991). Extended time for students with learning disabilities is a particularly difficult issue when trying to determine which accommodations are effective and valid. In addition, allowing students with learning disabilities extended time is controversial because students who are defined as having a learning disability exhibit low academic performance in school and lower performance on achievement tests than on ability tests. The most common rationale for the accommodation of extended time is that students with learning disabilities characteristically take more time to complete a variety of timed tasks than students without disabilities because of lower processing speeds (Zuriff, 2000). For instance, students with learning disabilities score significantly lower than students without disabilities under timed conditions on reading comprehension (Runyan, 1991a). However, Heaney and Pullin (1998) found that when students with disabilities were admitted on the basis of extended timed tests that the student’s first year college performance and future educational performance could be overpredicted by the test scores.

Because all students tend to benefit from extended time in test situations, it is important that time limits for all students are set to allow for optimal performance without unfairly advantaging one group over the other (Geisinger, 1994; Zuriff, 2000).

It is clear that practitioners must develop guidelines for deciding when ETTA is an appropriate testing accommodation and perhaps also assess the appropriate magnitude of the extended time that is appropriate for each individual (Willingham et al., 1988) as it is not clear that all students with disabilities
need the same amount of extended time to have valid test scores.

**How to Validate Test Accommodations via Differential Boost**

A well-controlled study of ETTA should compare standard and accommodated administrations of parallel test forms for students with and without disabilities (Tindal, 1998; c.f., Sireci, Li, & Scarpati, 2003). To validate the accommodation, ETTA must work with students with disabilities and must not work for students without disabilities (Tindal, Heath, Hollenbeck, & Harniss, 1998). Phillips (1994) argued that one important indicator that an accommodation serves to level the playing field between students with and without disabilities is the *differential boost* which is seen when an accommodation increases the performance of students with disabilities more than the accommodation increases the scores of students without disabilities.

Differential boost is an empirical manifestation that the accommodation speaks to something essential about the disability (Fuchs, 2001; Fuchs, Fuchs, Eaton, Hamlett, Binkley, et al., 2000). Most researchers defined differential boost as students with disabilities profiting substantially more than students without learning disabilities. To determine the numerical value of the differential boost, the researchers took the average boost for students without learning disabilities and added this average to one standard deviation of the nondisabled student’s boost (Fuchs, 2001, p. 177). Unfortunately, for students with learning disabilities, identifying accommodations that produce differential boost is not straightforward (Fuchs, Fuchs, Eaton, Hamlett, Binkley, et al., 2000).

**Maximum potential thesis.** Many of the empirical studies examining extended time as a test accommodation reference the Maximum Potential Thesis (MPT) which states that only students with learning disabilities benefit from extended examination time because students without learning disabilities would already be working at their maximum potential under timed conditions and would not benefit from extended time (Runyan & Smith, 1991; Zuriff, 2000).

In order for a study to confirm the MPT, the study would have to have two outcomes: (a) students with learning disabilities would score significantly higher under untimed than under timed testing conditions and (b) students without learning disabilities would not score significantly higher under untimed conditions.

Several studies have appeared to confirm the MPT hypothesis (e.g., Runyan, 1991a, b; Weaver, 1993). For example, Runyan (1991a) studied 31 students (16 students with learning disabilities; 15 students without learning disabilities) and found that under the timed condition, the comprehension scores for students without learning disabilities were significantly higher than those of the students with learning disabilities; however, the comprehension scores were not significantly different for the two groups under the untimed conditions. Therefore, the data supported the MPT in that students with learning disabilities benefited from extended time and reached their full potential, but students without learning disabilities did not benefit from extended time, presumably because they were already performing at their maximum potential under standard timed conditions. However, in many of these studies, students were not allowed to go back and change answers during the extended time. Students were only allowed to finish the test, thereby limiting the generalizability of these findings to practice (see also Zuriff, 2000).

A recent review of over 40 studies (Sireci, Li, and Scarpati, 2003) concluded that contrary to the MPT assumption, students without learning disabilities also benefit from ETTA, although not as much as the students with learning disabilities (for examples of some studies reporting these effects, see Halla, 1988; Hill, 1984; Zuriff, 2000). The studies conducted by Fuchs and her colleagues (Fuchs, Fuchs, Eaton, Hamlett, & Karns, 2000; Fuchs, Fuchs, Eaton, Hamlett, Binkley, et al., 2000) show that, on average, students with and without learning disabilities benefit comparably from many types of reading and math test accommodations. For these accommodations, there were no significant differences in the magnitude of benefit between students with and without identified learning disabilities, and effect sizes for students with and without disabilities were similar. Only 23% of students with learning disabilities benefited
substantially more than students without identified learning disabilities from accommodations on reading assessments (Fuchs, 2001, p. 177).

These findings raise serious questions about the selective use of ETTA for only students with disabilities and may point to more fundamental problems with testing practices, specifically whether current restrictions on test time are appropriate for students without identified disabilities.

**Implications for Research and Practice**

Extended time probably does allow students with identified disabilities to perform nearer their full potential than testing under more restrictive conditions. But to the extent that other students are not allowed extended time (and hence, to perform to their potential), ETTA may actually advantage students with learning disabilities (Fuchs, Fuchs, Eaton, Hamlett, Binkey, et al., 2000). This is an issue that transcends special education and learning disabilities, as the practice is prevalent, often unchallenged, and more often than not unjustified. The research above suggests that most of the time most students can benefit substantially from extended testing time. Well-controlled, valid research is needed (a) to demonstrate differential boost for ETTA (Fuchs, Fuchs, Eaton, Hamlett, Binkey, et al., 2000) and (b) to identify when ETTA is appropriate, and for whom, given that LD students are tremendously heterogeneous (Fuchs, 2001; Zuriff, 2000). Until the literature provides further guidance, practitioners may need to consider giving all students extended time on assessments to perform to their full potential.

Elliott et al. (2001) cautioned that in some cases providing extended time on a task may actually change the nature of the task being measured. To be sure, there are cases where authentic and valid assessment requires time constraints. Doctors training for emergency medicine careers, for example, must be able to make accurate diagnoses and decisions under severe time pressure in order to be effective. However, this rarely holds true for most academic courses of study. Does it really matter in a fundamental sense if a student takes an average of one minute per organic chemistry question or three minutes? Whether a student can complete a mathematical proof in five minutes or ten minutes?

That a student can write a coherent paragraph in a certain time period? Generally not. Instead of looking at ETTA as a special exception for special populations, the educational community in general should examine testing practices everywhere, for all students, and make timed tests the exception, not the rule, thereby leveling the playing field for all students (Mehrens, Millman, & Sackett as cited in Zuriff, 2000). Since untimed testing as a rule is impractical, researchers and practitioners need to examine ways to develop tests that will assess the construct of interest instead of the student’s speededness.

Future research should also explore the effects of assessment design and standardization to see whether incorporating assessment methodologies and techniques (i.e., computer mediated assessment, Rasch measurement methodologies, flexible testing conditions) reduces the need for accommodations while facilitating measurement of the critical constructs for all students (e.g., Thompson, Blount, & Thurlow, 2002).

There is also some anecdotal evidence that certain types of giftedness (e.g., visual-spatial giftedness) can actually impair test performance on timed tasks even when not paired with a learning disability. This is an area of accommodation research that has yet to be addressed, yet with gifted students accounting for a significant portion of the school-age population, this is a topic in dire need of study. Again, this reinforces the likelihood that ETTA is not just an important accommodation for students diagnosed with a learning disability.

In sum, current research regarding testing accommodations for students with disabilities highlights some of the challenges involved in determining how to fairly and validly assess students with disabilities and to provide appropriate accommodations. First, test scores must provide realistic, not inflated, estimates of a student’s ability to allow an accurate reflection of the student’s ability (Fuchs, Fuchs, Eaton, Hamlett, Binkley, et al., 2000). Second, test administrators must determine if test instruments are suitable for use with students with disabilities (Geisinger & Carlson, 1995). Third, they should consider whether students with learning disabilities were included in the normative and validation samples (Geisinger & Carlson, 1995).
Fourth, practitioners need to know that when published instruments are adapted that interpretations of the results should be tentative. The modified assessment simply is not the same measure as the original version for which norms and validation results exist (Geisinger & Carlson, 1995). And finally, practitioners must weigh whether the assessment without validation is going to yield useful information over and above which is already available from non-test sources. Ultimately, practitioners must determine if a nonstandard assessment, such as one with ETTA, is any less valid than a standard administration of the same assessment. This is also an important question for future research.

**How No Child Left Behind Makes ETTA a Critical Issue**

Under NCLB, standardized test scores are used to determine schools in need of improvement and explicitly demands that most students participate in high-stakes testing. The combination of tests that determine important outcomes and the demand that most of the school population participate combines to make the issue of accommodation and validity of test scores critical.

Unfortunately, what we know about ETTA and other accommodations thus far is that optimal performance for all students would probably come if many common LD accommodations were implemented for all students—extended or flexible time, distraction-free test settings, and so on; therefore, allowing all students to perform to their potential, giving schools the best assessment of student learning possible. At this point, we do not appear to have the science to either (a) reliably discriminate between students who require ETTA and those who can perform to their potential under standard time allowed, or (b) exactly what the optimal time allowed for all students should be.

**REFERENCES**


Department of Justice (1996). Nondiscrimination on the basis of disability by public accommodations and in commercial facilities rule, 28 C.F.R.§36.309.


Stretch & Osborne, Extended Test Time Accommodation


**Citation**


**Authors**

Dr. LoriAnn S. Stretch is an adjunct faculty member in the School of Education at Campbell University in Buies Creek, NC, and Dr. Jason Osborne is an Associate Professor of Educational Psychology at North Carolina State University at Raleigh. Correspondence regarding this article should be sent to LoriAnn S. Stretch at 216 Alder Lane, Clayton, NC 27520 or LStretch@ClaytonCounseling.com