How Much Time?: A Review of the Literature on Extended Test Time for Postsecondary Students with Learning Disabilities

Nicole S. Ofiesh
University of Arizona

Charles A. Hughes
The Pennsylvania State University

Abstract

One ongoing dilemma with the accommodation of extended test time is how much time to provide. Due to a dearth of research to help disability service providers with this decision, a review of the literature on extended test time for postsecondary students with learning disabilities (LD) was conducted to (a) inform service providers about the results of several studies on extended test time, (b) determine if a certain amount of extended test time was typically used by participants with LD, and (c) identify research variables from the studies that could account for differences in the amounts of time use. A search resulted in seven studies that included reports of time use. The average time use in most studies ranged from time and one-half to double time. Differences in results based on type of postsecondary setting, test conditions and test instruments are discussed, and recommendations are offered to guide the decision-making process on how much additional time to provide.

The right to test accommodations in postsecondary educational settings stems from regulations accompanying statutory law (e.g., Americans with Disabilities Act [ADA], 1990; Section 504 of the Rehabilitation Act of 1973). Of the various ways to accommodate students with learning disabilities (LD), extended test time is the most frequently requested and granted in colleges and universities (Bursuck, Rose, Cohen, & Yahaya, 1989; Nelson & Lignugaris-Kraft, 1989; Yost, Shaw, Cullen, & Bigaj, 1994).

The accommodation of extended test time is built on a growing body of literature that supports the contention that individuals with LD characteristically take longer to complete timed tasks, including taking tests (e.g., reading passages, math calculations), than individuals without these disabilities (Bell & Perfetti, 1994; Benedetto-Nash & Tannock, 1999; Chabot, Zehr, Prinzo, & Petros, 1984; Frauenheim & Heckerl, 1983; Geary & Brown, 1990; Hayes, Hynd, & Wisenbaker, 1986; Hughes & Smith, 1990; Wolff, Michel, Ovrut, & Drake, 1990). This slowed rate of performance prevents some students with LD from completing as much of a test as their peers, leading to lower scores. When provided with additional time, however, many students with LD are able to finish more of a test and thereby make significant gains in their test score (Alster, 1997; Hill, 1984; Jarvis, 1996; Ofiesh, 2000; Runyan, 1991a, 1991b; Weaver, 2000).

Conversely, extended time often does not benefit students without LD in the same way. On the majority of tests used in studies to assess the effectiveness of extended test time, students without LD as a group either (a) did not use the extra time, or (b) did not make significant score gains with the use of more time (Alster, 1997; Hill, 1984;

Ofiesh, 2000; Runyan, 1991a). However, because some students without LD may demonstrate score increases with extended test time, it is important to clarify that the purpose of a test accommodation is to ameliorate the difference between individuals with and without disabilities. A test accommodation like extended test time should not accommodate nondisability-related factors that can impact test taking for all students (fatigue, test anxiety, motivation, test-taking skills). Thus an important question becomes, "How much extra time is reasonable and fair?" Too little time will not accommodate the disability. Too much time may accommodate the disability, as well as nondisability-related factors such as motivation or anxiety, and therefore provide an unfair advantage to students without LD. Furthermore, for the student with LD, too much time may result in test scores that are an invalid representation of academic ability or achievement (Braun, Ragosta, & Kaplan, 1986; Ziomek & Andrews, 1996; Zurcher & Bryant, 2001).

In practice, the process of deciding how far to extend the time limit of a test is not clearly defined, and in most instances there is no precise answer. Rather, postsecondary disability service providers (DSP) estimate an amount of time based on a variety of factors such as the disability services program policies, the individual student's strengths and weaknesses, the test, the program of study, and other unique information (e.g., previous history of accommodation, medication). However, new studies exist to assist DSP on how to weigh these factors and where to begin with this decision, with respect to the ADA.

The goals of this article are twofold. The first is to provide individuals who are responsible for determining appropriate accommodations with a review and analysis of the literature on extended test time with respect to time use. Such information also provides researchers with a foundation for further investigation into the accommodation of extended test time. The second goal is to provide DSP with a benchmark (i.e., a starting point) from which to gauge their decisions about extended test time. To accomplish these two goals, the literature was analyzed to determine if a certain amount of extended test time was typically used by participants with LD in studies on extended test time. Furthermore, research variables were identified that could account for differences in the amounts of time use among the studies (e.g., type of postsecondary institution or type of test) (Runyan, 1991b). The article begins with an introduction to how extended test time decisions are made in postsecondary settings, followed by a review and analysis of the studies, discussion and recommendations.

Determining Appropriateness of Extended Test Time

It is usually the role of the disability service provider and/or the ADA coordinator to determine the reasonableness of a student's request for an accommodation based on a disability, in relation to precepts from the ADA. These precepts are (a) the current impact of the disability on a major life activity, and (b) the functional limitations of the disability. This information about an individual's disability is, in part, documented in the student's written diagnostic evaluation. Recent survey research has indicated that most DSP use a student's diagnostic evaluation to help make decisions about service delivery, including accommodations (Ofiesh & McAfee, 2000).

In the same research by Ofiesh and McAfee (2000), DSP ranked the most useful section of the written evaluation to be the diagnostician's summary of the student's cognitive strengths and weaknesses. This section often details the functional limitations of a student with LD and therefore helps to determine the reasonableness of an accommodation request. Even so, while DSP rated this section most *useful*, they reported that in the end, they most often *used* (i.e., relied upon) the diagnostician's professional recommendations for their service delivery decisions. Additionally, some respondents noted that the sole use of diagnostic evaluations to make service delivery decisions was ineffective because frequently other "potentially useful information" such as history of accommodation use, current impact of disability on the different academic areas, and other exceptional conditions was missing. The need for more informa-

tion to make sound accommodations decisions is not unique to DSP and the type of information needed is like that used in the accommodation decision-making process by national testing agencies (Educational Testing Services, 1998; National Board of Medical Examiners, 2001). In practice, some DSP reported that they gather the necessary information through interviews and informal assessments of their students to supplement the diagnostic evaluation. However, in determining extended test time accommodations, DSP must also consider the characteristics of the specific test to be accommodated. While some diagnosticians relate functional limitations to certain types of tests, others do not make this connection. In some instances it is simply not practical for a diagnostician to detail the functional limitations of an individual's disability in terms of all the types of course tests a student may encounter and need accommodated (e.g., essay, only math, all multiple choice tests/all subjects). Thus, DSP commonly make their own inferences about functional limitations as they relate to specific course tests.

Two important considerations include the type of test (e.g., essay) and the informational content on which the student is being tested (e.g., reading comprehension, calculus). If time is *not* an essential component of the test (e.g., a test of factual knowledge) and a student's disability significantly impacts the ability to demonstrate what he or she knows and can do under timed circumstances, the student may qualify for extended test time. This is the most typical scenario in postsecondary settings. However, there are other instances where time may be an essential component of a course test (e.g., a timed sprint in a physical education class) or the instructor may desire to make speed a component of a test (e.g., a 5-minute pop quiz, a firm one-hour medical ethics test). In these cases, the course instructor and the person responsible for authorizing accommodations must determine if extended time will invalidate a test, or remove an essential component from the course or a program of study. On occasion, the discussion requires mediation at a higher administrative or legal level. Most important, the DSP must make test accommodation decisions that maintain the validity of the test based on its purposes, and the specific inferences made from test scores (Wainer & Braun, 1988). Once extended test time is determined to be appropriate for a certain individual, DSP are left with the determination of how much time is appropriate.

Gauging How Much Time

Anecdotal data suggest that practice varies throughout offices for disability services regarding how to gauge the amount of extended test time a student may need. Both conservative and liberal timing can be found in current practice. For example, some DSP rely on one standard amount of time for most, others use ranges from 25%-400% extended time and, though rarely, others provide unlimited time. One approach to gauging the amount of time, as recommended by professionals in the field and in the literature (Alster, 1997; Fink, Lissner & Rose, 1999; Jarvis, 1996; Ofiesh, 1999; Ofiesh, Brinkerhoff, & Banerjee, 2001; Runyan, 1991b; Weaver, 1993), is to synthesize a variety of information about the student, test and program of study, and evaluate a preponderance of evidence for each request individually. However, empirical research on the factors that most relate to the need for, and influence the amount of, more time is still at its early stages (Ofiesh, 2000; Ofiesh, Kroeger, & Funckes, 2002), and limited data are available to assist DSP in knowing how to weigh certain factors in the synthesis of information.

Some individuals have begun to systematically collect data at their own institutions in order to have a better understanding of how certain variables influence how much time is reasonable and fair. For example, service providers at the University of California at Los Angeles (UCLA) found one way to consider factors related to test characteristics and program demands at their institution. At this university, 13 subject areas were evaluated by the amount of extended time used by students with LD. Considerable differences were noted among academic areas and the practitioners suggested that DSP could gauge the amount of time a student needed, *in part*, by evaluating similar data at their own institutions ("Use Research,"

2000). In the meantime, there clearly appears to be a desire on the part of DSP to be well informed and to make defensible decisions in a professional, ethical, legal, and empirically based manner. It is our intent through this article to disseminate research-based recommendations to promote this worthwhile practice.

Method

A computer search was conducted using the search engine Silver Platter, with the databases Educational Resources Information Center (ERIC) and Dissertation Abstracts International (DAI), to identify studies investigating extended test time for postsecondary students with LD. The search terms included, "extended test time," "test accommodations," "accommodations," and "testing" <and> "students with disabilities." It was predetermined that all dissertations and empirical studies published in refereed journals between 1980-2001 on the subject of extended test time for postsecondary students with disabilities would be (a) included for consideration in the review,

Table 1

Studies on the Effectiveness of Extended Test Time for Adults with LD

Alster, E. H. (1997). The effects of extended time on the algebra test scores for college students with and without learning disabilities. *Journal of Learning Disabilities, 30,* 222-227.

Halla, J. W. (1988). A psychological study of psychometric differences in Graduate Record Examinations General Test scores between learning disabled and non-learning disabled adults (Doctoral dissertation, Texas Tech University, 1988). *Dissertation Abstracts International, 49,* 194.

Hill, G. A. (1984). Learning disabled college students: The assessment of academic aptitude. (Doctoral dissertation, Texas Tech University, 1984). *Dissertation Abstracts International, 46,* 147.

Jarvis, K. A. (1996). Leveling the playing field: A comparison of scores of college students with and without learning disabilities on classroom tests (Doctoral dissertation, The Florida State University, 1996). *Dissertation Abstracts International*, *57,* 111.

Ofiesh, N. S. (1997). Using processing speed tests to predict the benefit of extended test time for university students with learning disabilities. (Doctoral dissertation, The Pennsylvania State University, 1997). *Dissertation Abstracts International, 58,* 76.

Ofiesh, N. S. (2000). Using processing speed tests to predict the benefit of extended test time for university students with learning disabilities. *Journal of Postsecondary Education and Disability, 14*, 39-56.

Runyan, M. K. (1991a). The effect of extra time on reading comprehension scores for university students with and without learning disabilities. *Journal of Learning Disabilities, 24,* 104-108.

Runyan, M. K. (1991b). Reading comprehension performance of learning disabled and non learning disabled college and university students under timed and untimed conditions (Doctoral dissertation, University of California, Berkeley, 1991). *Dissertation Abstracts International, 52,* 118.

Weaver, S. M. (1993). The validity of the use of extended and untimed testing for postsecondary students with learning disabilities (Doctoral dissertation, University of Toronto, Toronto, Canada, 1993). *Dissertation Abstracts International, 55,* 183.

Weaver, S. M. (2000). The efficacy of extended time on tests for postsecondary students with learning disabilities. *Learning Disabilities: A Multidisciplinary Journal*, *10*, 47-55.

Note. Runyan's 1991a study was the pilot research for her dissertation (1991b); therefore Runyan 1991a and 1991b are not the same study.

and (b) analyzed to determine if the results presented data on the participants' use of extended test time. Only those studies that reported the amount of time used under extended test time conditions for students with LD were included for purposes of this investigation.

No studies were located that specifically addressed the issue of "how much time" postsecondary students with LD used. Ten studies were identified in which the effectiveness of extended test time for postsecondary students with LD was investigated (see Table 1). Seven reported amount of time used and were included in the literature review for analysis. When amounts of time were not reported, the data needed for this investigation could not be acquired, and these studies consequently were not included in the review (Ofiesh, 2000; Runyan, 1991b; Weaver, 2000).

Analysis of Selected Studies

Each study was analyzed to identify (a) the dependent variable (i.e., test instruments), the independent variables or conditions that provided the participants with more time (e.g., standard, extended, unlimited), (c) the standard test administration time, (d) the participants' range of total test time with extended time conditions, and (e) the average amount of extended time participants used, in relation to the standard administration time. Once the amount of participants' total test time use was determined through either a reported mean (e.g., average of 25 minutes for the group to complete the test) or a range of performance (e.g., 21-32 minutes for the group to complete the test), the average amount of extended time was calculated for each dependent variable.

To determine the average amount of extended time needed to complete a test, the mean amount of extended time for the group was divided by the standard test administration time. For example, in one study (Alster, 1997), the standard test administration time was 12 minutes. Under the extended test time condition, students with LD took 25 minutes to complete the test. Dividing the mean time use (e.g., 25 minutes) by the standard administration time (e.g., 12 minutes), the result 2.1 indicated that students with LD in that study took approximately double time to complete the test. In two of the seven studies, a range was reported without a mean (Jarvis, 1996; Runyan, 1991a). In these cases, the mean was calculated based on the midpoint of the range and should be interpreted with caution. The Results section presents the stated purpose(s) and findings of each study. Important variables that influenced the outcomes of the studies are presented as each study is discussed, followed by a separate section on time use.

Results

Summary of Studies Reporting Additional Time Use Under Extended Test Time Conditions

Seven studies identified the actual amount of time participants used under extended test time conditions. A summary of the test instruments, test conditions and the standard, mean, and additional amounts of time study participants used is presented in Table 2. All studies employed quasi-experimental designs and included students with and without LD, attending postsecondary institutions.

The studies included a variety of tests to measure the impact of varying time conditions on test performance. Tests included (a) Nelson-Denny Reading Test (NDRT) (Brown, Bennett, & Hanna, 1981; Brown, Fishco, & Hanna, 1993), either as a total score or one or both subtests (i.e., Vocabulary and Comprehension); (b) ASSET Elementary Algebra Test (American College Testing Program, 1989); (c) American College Test (ACT) Social Studies, English, and Math tests (American College Testing Program, 1981); (d) Graduate Record Examination (GRE) (Educational Testing Service, 1986); and (e) actual classroom tests (Jarvis, 1996); all under a variety of time conditions.

Table 2 denotes the independent variable or condition with the exact titles the researchers used to label the variables or conditions in their studies (e.g., "unlimited time"). However, since the meanings of the labels were used inconsistently among the researchers, the operational definition of each condition is also noted. For example, Alster (1997), Runyan (1991a), Jarvis (1996), and Weaver (1993) used the terms "extended time" and "extra time" to describe a condition where

Table 2

Time Usage of Participants with LD Under Additional Time Test Conditions

Author	Participants	Dependent Variable (standard time administration in hours/minutes)	Independent Variable (test time condition)	Range and mean of time use with more time (in hours/minutes) for students with LD	Time use under additional time condition divided by standard time
Alster, E. (1997)	N=88 LD n=44	ACT ASSET Algebra Test (12 m)	Timed (standard) Extended time 1	 (11-56m), x=25m	 2.1
Hill, G. (1984)	N=96 LD n=48	ACT Social Studies, English, Math (2h, 40m) NDRT (1981) Total Score (35 m)	Timed (standard) and Untimed 2	ACT x=4h, 4m NDRT x=1h, 14m	1.5 2.1
Halla, J. (1988)	N=126 LD n=66	GRE (3h, 30m) NDRT (1986) Total Score (35m)	Timed (standard) Untimed 3	 GRE x=3h, 17m NDRT x=50	 0.9 1.4
Jarvis, K. (1996)	N=157 LD n=40	Classroom Tests (50 m)	Timed (standard) Test 1 and 2		
			Extended time 1 Test 3 and 4	Test 3 x=1h, 15m Test 4 x=1h, 11m	1.4 1.4
Ofiesh, N. (1997)	N=60 LD n=30	NDRT (1993) Total Score (35 m)	Timed (standard) Extended time 4	 NDRT Total x=45m	 1.3
Runyan, M. (1991a)	N=31 LD n=16	NDRT (1981) Comp. (20 m)	Timed (standard) Extra time 1	 NDRT Comp. (24-49m) x=36	 1.8
Weaver, S. (1993)	N=88 University students with LD n=31	NDRT (1981) Voc. (15) Comp. (20 m)	Timed (standard) Extended time 1	 Extended time Uni Voc x=22 m Col Voc x=32 m	 Extended time Uni Voc 1.5 Col Voc 2.1
	College students with LD n=8			Uni Comp x=27 m Col Comp x=38 m	Uni Comp 1.4 Col Comp 1.9
			Untimed 5	Untimed Uni Voc x=31 m Col Voc x=35 m	Untimed Uni Voc 2.0 Col Voc 2.3
				Uni Comp x=35 m Col Comp x=34 m	Uni Comp 1.8 Col Comp 1.7

1 Participants were first given the standard amount of time, then when time was up they were told to take as much time as needed to finish the test.
2 Participants were explicitly told to take as much time as needed and to take the test over more than one session if necessary.
3 Participants were given several tests at once and told to finish as much as they could during the additional test time, then to finish over as many additional sessions as needed.
4 60% more time than standard; the students were told how much time they would have for each test: 24 m for Vocabulary and 32 m for Comprehension.
5 Participants wrote in a room by themselves and were told they could have all the time they needed to finish the test.
Note. When time usage was reported in ranges, means were calculated by the midpoint of the range (Jarvis, 1996; Runyan, 1991a).

participants were allowed to take as much time as needed to finish the test once the standard administration time was up. Ofiesh (1997), on the other hand, used the term "extended time" to describe a condition where participants were given 60% more time than standard on an alternate form, and the students were told at the beginning of the test how much time they would have.

One of the first controlled studies to assess the effects of untimed testing conditions on the validity of academic and ability tests for students with and without LD was conducted by Hill (1984), who evaluated the impact of timed and untimed testing on test scores, and the relationship of those scores to grade point average (GPA). For the participants with LD, all three ACT tests and the two NDRT subtest mean scores were higher in the untimed testing condition than in the timed testing condition. However, for the participants without LD, the Vocabulary subtest of the NDRT was the only subtest for which the mean score was significantly higher in the untimed testing condition than in the timed testing condition. Furthermore, Hill found no differences between the correlations of timed or untimed ACT test performance and GPA, concluding that the untimed ACT score was a valid predictor of college GPA for students with LD only. Students without LD correlated with GPA only under standard time conditions.

In terms of time usage and test completion, Hill found that the percentage of completed test items for students with LD under untimed conditions was nearly 100%, but substantially lower with set time limits. Since participants were allowed to take as much time as desired, it is not clear why all students with LD did not complete 100% of the test under untimed conditions. It is possible that some did not want to guess, a practice that is commonly recommended on some standardized tests. However, for the participants without LD the percentage of items completed did not change with more time. When given unlimited time, the average amount of time use for students without LD on the ACT and NDRT was less than for students with LD, amounting to 3 hours and 5 minutes on the ACT tests and 1 hour on the NDRT.

Halla (1988) used the NDRT and the GRE to study the effects of extended test time on score performance for students with and without LD. Her basic results diverged significantly from Hill's and those of subsequent researchers by the finding that students with and without LD showed no difference in timed scores. Both students with and without LD made substantial score gains under an unlimited time condition, even though students with LD, on the average, did not use the extra time. Furthermore, the students without LD used approximately 21 minutes more on the GRE than students with LD, and both groups used the same amount of time on the NDRT.

Two factors may have confounded the outcome of this study. First, there was a significant difference between intelligence scores (IQ) of the participants with and without LD. The average IQ for participants with LD was 120.86 and the average IQ for students without LD was 111.91. Halla noted that when a secondary analysis controlled for IQ, the results changed. In the groups of students with and without LD whose IQs were 117 and below, participants with LD scored significantly lower than students without LD under timed conditions. Moreover, students with LD made enough gains under unlimited time to perform at par with their nondisabled peers. A second confounding variable could be that the participants were told that the purpose of the study was to assess variable time conditions on performance, thus possibly influencing their performance on the exact variable being measured. Since the Hill and Halla studies conflicted so dramatically, Runyan's study helped to clarify previous findings.

Participants in Runyan's study (1991a) were students with and without LD from the University of California at Berkeley. Results clearly demonstrated that students with LD made greater score gain than students without LD under extended test time conditions on the Comprehension section of the NDRT. Furthermore, the scores of students with LD under the extended time condition were commensurate with both the standard and the extended-time scores of students without LD. Runyan controlled for ability using SAT

scores, and the findings paralleled Hill's on the NDRT in terms of the need for more time among students with LD only. In terms of time use, the students with LD all used more time to finish the test, but only two of the students without LD needed more time. These two students finished the test with 3 - 4 minutes more.

Weaver (1993, 2000) confirmed the findings of Hill and Runyan for students with and without LD and added a condition where the student was tested privately with the test untimed. While both students with and without LD made some score gains under extended and untimed conditions, only students with LD made significantly greater gains than students without LD. Unlike previous researchers, Weaver hypothesized and confirmed that there would be significant differences in test performance (i.e., amount of gain and time use) between students from different types of postsecondary institutions under varying time conditions. To test this hypothesis, she included college students with and without LD (i.e., students from an open admissions school) and university students with and without LD (i.e., students from a competitive school). Like in the Runyan (1991a) study, students without LD needed little more than 1 - 4 minutes to complete the NDRT, but students with LD needed and benefited from more time (see Table 2). Because the Hill, Runyan, and Weaver studies had similar findings, subsequent investigations were designed to evaluate new aspects of the extended test time question. These included actual classroom tests, math tests, and the use of speeded diagnostic tests to predict the benefit of extended test time.

Jarvis (1996) studied the effects of extended test time on four combined short-answer and multiple-choice actual classroom tests at Florida State University. Her results diverged from all previous findings and the implications are not clear. Specifically, the performance of students with LD under extended test time was similar to that of students without LD under standard time. However, the difference between standard and extended test time was not significant for students with LD, but was significant for students without LD. Additionally, students without LD used, on the average, only 1 - 5 minutes more than students with LD. Jarvis attributed her performance findings for the groups of students with and without LD to low statistical power, a consequence of small sample sizes in the control and treatment groups. Another important consideration is that students with and without LD self-selected to participate in the extended time condition. Although the sampling procedure made an attempt to randomize, the treatment was self-selected. For both students with and without LD, it is likely that the students who elected to participate in the extended time conditions were ones who assumed they would benefit, or the results would have changed if a greater number of students would have selected the option.

Alster (1997) examined the effects of extended time on the algebra test scores of community college students with and without LD. Findings supported previous research in that students with LD made significant score gains with extended test time, whereas their peers without LD did not (Hill, 1984; Runyan, 1991a; Weaver, 2000), even though the students without LD spent an average of 20 minutes on the 12-minute test when given extended time. This was only 5 minutes less than the average amount of time students with LD spent on the test when given more time.

Building on the growing body of literature favoring significant performance differences between students with and without LD under extended test time, Ofiesh (1997) investigated the validity of the relationship between diagnostic tests of processing speed and extended test time for students with and without LD. Using the NDRT total test score, a significant relationship was found between processing speed and the benefit of extended test time for students with LD only. Ofiesh's study differed from previous studies on extended test time in that she controlled the amount of extra time participants were given—slightly more than time and one-half. Furthermore, she notified students of the amount of time in both the standard and the extended-time administrations and used alternate forms for the conditions instead of telling participants to complete the test when the standard time was up. Under these conditions,

previous findings on test performance under extended-time conditions between participants with and without LD were supported, although the amount of time needed to finish the test was considerably less.

Two reasons could have accounted for this difference. First, students may allocate time and approach a test differently when told how much time will be allowed. Second, Ofiesh used a newer version of the NDRT than previous researchers had used. In 1993 the Vocabulary section of the NDRT was shortened from 100 to 80 items, but the administration time remained the same. The newer slightly modified version reduced the completion rate for test takers in the normative sample from 6.7 items per minute to 5.3 items per minute. Furthermore, in the Comprehension section, the number of selections was changed from eight to seven shorter selections, but with five instead of four questions for each section (Brown, Fishco, & Hanna, 1993).

Average Amounts of Extended Time

In most studies where students were instructed to take as much time as they needed to finish, they usually used an average of more than time and a half but not much more than double time (e.g., 2.1). The exception was the performance of university students with LD in the Weaver study on the Comprehension section of the NDRT ($\underline{M}$ = 1.4) and in the Ofiesh study on both sections of the NDRT ($\underline{M}$ =1.3). Since the ranges of time use were reported in four of the studies (Alster, 1997; Jarvis, 1996; Ofiesh, 1997; Runyan, 1991a), it was possible to determine the highest and lowest possible amount of time usage. The largest range was found on the ASSET, where at least one individual with LD used quadruple time to complete the ASSET and at least one individual completed the test 1 minute under standard time (Alster, 1997).

Discussion

Contributions of the Studies to the Determination of How Much Time

Time and one-half to double time as a general rule. The results of the analysis of time use suggest that the range of time and a half to double time as a basis for decision making is a good place to start and provides enough time for most students with LD to finish a test (Alster, 1997; Hill, 1984; Jarvis, 1996; Ofiesh, 1997; Runyan, 1991a; Weaver, 1993).

It is important to keep in mind that accommodation decisions must be made on an individual basis. The averages of time use from the studies are fairly consistent, especially on the 1981 version of the NDRT; yet amounts of time are *averages* based on aggregated data and mean performance times, not individual performance. Some individuals used no extra time, less than time and one-half; and some used more than double time, though less frequently.

Double time may be liberal for some students. For example, the study by Ofiesh (1997) suggested that students with LD might take more time than needed when given that additional time, simply because they were given more time. Moreover, the averages that are close to double time may have been a result of the tightly timed nature of the standardized tests used in those studies. While most classroom tests are designed to be finished by all students, standardized tests often are not designed to be finished by all test takers. For example, Alster noted that the reported completion rate of the ASSET is 69% (Alster, 1997). Therefore, close to 30% of the students would not be expected to finish the ASSET in the allotted standard time. While it can be concluded that students with LD needed additional time to finish a test, the use of double time may have been influenced by the test's built-in completion rate. In other words, if a test cannot be completed by all test takers, then it may take an unusually greater amount of time to finish than a test designed to be finished by most (e.g., a classroom test). However, in support of the need for double time on the ASSET, the summary of data collected at UCLA

("Use Research," 2000), ranked math as the fourth highest area among the academic subjects needing the largest amount of additional time for students with LD.

This analysis can help frame disability service policies at college campuses. At the very least, it is important to be clear about what additional time options a program or office for students with LD provides (e.g., 25%, 50%, 100%, unlimited). Unlimited time in postsecondary settings is not common practice, but some psychoeducational diagnosticians routinely recommend this accommodation. Clarity about time options would help resolve problems with uncommon recommendations from diagnosticians and student requests before they are presented.

Differences among postsecondary institutions. Weaver (1993) suggested that differences in the amount of additional time used by students with LD vary significantly with the type of postsecondary institution. Runyan (1991b) also stated this hypothesis. Both researchers speculated that this could be due to the characteristics associated with an institution's admissions requirements and subsequent student body. While Weaver compared two types of institutions, one with open admission and one with competitive admission, Runyan compared students from a competitive four-year institution with students from a community college. In both studies the students in differing institutions demonstrated significant differences in the amount of time it took to complete the test. Since the average intelligence of a student body can change as a function of admissions requirements (Longstreth, Walsh, Alcorn, Szeszulski, & Manis, 1986), these findings also relate to Halla's conclusion that the IQ of students with LD can impact performance under differently timed conditions**.** One way to address the heterogeneity in test performance among student populations is to analyze the test-taking performance and accommodation decision process for students at institutions separately (Ofiesh, 2000). Known as local norms, service providers at postsecondary institutions have used this type of data effectively to evaluate the specific characteristics of students with LD within a specific student body (Mellard, 1990). Therefore, service providers working with students with LD are encouraged to begin to collect a database from their own student body (i.e., local norms) for all students who receive extended time via a simple coding sheet. Important data to collect include (a) amount of test time provided, (b) amount of test time used, (c) amount of time used by subject area, (d) typical length and format of exams by instructor (e.g., essay, case study), (e) selected diagnostic tests and student percentile or standard scores, and (e) diagnostician's recommendations for time (Ofiesh, Hughes, & Scott, 2002). Such information would allow DSP to begin to make decisions regarding the amount of time to provide in a systematic way, grounded in their own data and expert judgment. Ultimately, practitioners will be able to evaluate and reflect on their approach in the decision of how much time to provide.

Use of indicators in psychoeducational evaluations. To begin to validate the use of psychoeducational documentation in determining when to give extended test time and how much time to give, one study investigated the correlation between processing speed test scores and the benefit of extended test time (Ofiesh, 1997). Results showed that the lower a person's processing speed score the greater the likelihood of benefit from extended test time for a student with LD. Replications of this study and new research investigating a more concrete relationship between these and other scores that impact speed (e.g., memory retrieval) and amount of time are in process; however, these findings suggest DSP can use the Cross Out and Visual Matching test of the WJ-R and the Reading Fluency test of the WJ-III as one indicator to gauge the need for and amount of extended test time (Ofiesh, 2000; Ofiesh, Kroeger, Funckes, 2002).

When considered, cognitive and academic tests reported in diagnostic evaluations should be interpreted in terms of standard scores and percentile ranks as these scores are measures of relative standing and can illustrate how an individual compares to peers in terms of speed of processing information. For example, when an individual receives a percentile score of 10 on a processing

speed or memory retrieval test, this means that 90% of the norm group performed that task faster and/or more accurately, regardless of IQ. Using the normal curve, a percentile of approximately 9 or lower is generally referred to as a "very low" score, and individuals, obtaining such scores may be the ones who need more than time and one-half. When several scores from selected constructs are evaluated in this manner, it allows a DSP to get a better idea of a student's speed of processing information in relation to a peer group. When used with local normative data (i.e., specific test data collected on the home university population), DSP can begin to draw inferences regarding what it means for their own students in terms of the need for additional time when a student falls in certain percentile ranges.

In order to make valid use of tests, it is useful to have an established list of acceptable types of cognitive, neuropsychological, intelligent and academic test and subtests as indicators for the need of more time. Tests that hold a reliability coefficient of .80 or higher are considered acceptable as screening tests for problem areas (Salvia & Ysseldyke, 2001) and could be used as part of evidence in the decision making process. In addition to test scores and summaries of strengths and weaknesses, a diagnostician's analysis or observation of test behavior can provide support to the data. Such qualitative data can serve as a supportive rather than primary source of information.

While documentation is important, it cannot be overstated that the review of psychoeducational or diagnostic evaluations should include a search for more than one indicator to gauge the need for time. That is, decisions should not be based on one test score, observation, or piece of data; an accumulation of data should be weighed. Once a profile based on test results and observations is established, other information such as the format of the test and the impact of the functional limitations of the individual on one or more academic areas must be factored in to gauge timing decisions.

Additional considerations to gauge amount of time. After data are evaluated in a holistic manner, the characteristics of the timed test must be considered. These include the (a) content (e.g., math, history), (b) format of the test (e.g., multiple-choice, short-answer, essay, combination), and (c) type of response (e.g., calculation, written, pictorial). The length of extended time may change with the test content.

Test time can also change as a result of test format, type of response required and the functional limitations of disability. If the test requires an essay response, for example, and many indicators suggest significantly low performance in visual-motor tasks and writing tests, an individual may need a greater amount of time than typically provided for a multiple choice test. Furthermore, other accommodations not related to test taking per se can add to or reduce the amount of time a student needs during a test situation (e.g., scribe, taped text).

Conclusion

In summation, this literature review provides professionals in disability services with an understanding of the typical amounts of extended test time students with LD used in several studies. This information, as well as other important variables related to these studies, has been presented in an effort to encourage effective decision making regarding extended test time for postsecondary students with LD. Disability service providers can use this knowledge as a benchmark or starting point for gauging the amount of time students may need to perform equitably with their nondisabled peers. Deciding how much additional time to grant is multifaceted and includes: (a) awareness of the average amount of time students use and an understanding that this amount of time can vary based on the postsecondary institution, (b) information from diagnostic tests in targeted areas, (c) an understanding of the classroom test characteristics, (d) an individual's functional limitations, and (e) an individual's characteristics (e.g., dual diagnoses, medication). With this information, significant evidence can be accumulated in order to decide how much time to grant for each student who is entitled to the accommodation of extended test time. The collection of local data based on the

characteristics of individual postsecondary institutions is highly recommended. Findings suggest that test performance and the need for more time can vary among institutions of higher learning. Using local data and the recommendations provided herein, DSP can begin to make decisions regarding test time that factor in the unique characteristics associated with their own student body.

Limitations

It is important to note that the studies in this literature review included postsecondary students with LD. Therefore, the recommendations should not be generalized to elementary or secondary students with LD. The heterogeneity of the LD population is too great to apply the conclusions of the current review beyond the postsecondary arena, and research on the effectiveness of extended test time for younger students is not clear (Fuchs & Fuchs, 2001; Munger & Loyd, 1991). The recommendations developed for DSP in this article should not be applied to practice in special test administrations of standardized tests such as college boards and graduate exams. In these cases timing decisions with the goal of granting enough time for all students with LD to *finish* the test—as is usually the situation on classroom tests—may not be equitable to students without disabilities. Ragosta and Wendler (1992) state, "the mean amount of time needed is not the appropriate measure to use when establishing the amount of time allowed for special test administrations" (p. 4).

Future Directions

More studies are needed to evaluate the use of time on actual classroom tests and the process used to make timing decisions. Additionally, studies are needed to clarify the factors that influence an individual's need for certain amounts of time. Further investigations into the validity of achievement and cognitive test scores found in diagnostic evaluations, and the decisions based on these scores, are also needed in order to validate this practice. Other emerging issues that need to be addressed in the arena of extended test time include the legitimacy of accommodating a test-taking strategy versus a disability with more time and the impact of psychiatric disabilities and medication on test time.

References

Alster, E. H. (1997). The effects of extended time on the algebra test scores for college students with and without learning disabilities. *Journal of Learning Disabilities, 30,* 222-227.

American College Testing Program. (1981*). ACT assessment supervisor's manual of instructions for special testing*. Iowa City, IA: American College Testing.

American College Testing Program. (1989). *ASSET technical manual*. Iowa City, IA: American College Testing.

Americans with Disabilities Act of 1990, 22 U.S.C.A. § 12101 *et seq.* (West 1993).

Bell, L. C., & Perfetti, C. A. (1994). Reading skill: some adult comparisons. *Journal of Educational Psychology, 86*, 244-255.

Benedetto-Nash, E. & Tannock, R. (1999). Math computation performance and error patterns of children with attention-deficit hyperactivity disorder. *Journal of Attention Disorders, 3,* 121-134.

Braun, H., Ragosta, M., & Kaplan, B. (1986*). The predictive validity of the GRE General Test for disabled students* (Studies of Admissions Testing and Handicapped People, Report No. 10). New York: College Entrance Examination Board.

Brown, J. I., Bennett, J. M., & Hanna, G. (1981). *Nelson-Denny Reading Test.* Boston, MA: Riverside.

Brown, J. I., Fishco, V. V., & Hanna, G. (1993). *Nelson-Denny Reading Test.* Chicago: Riverside.

Brown, J. I., Fishco, V. V., & Hanna, G. (1993). *Nelson-Denny Reading Test manual for scoring and interpretation.* Chicago: Riverside.

Bursuck, W. D., Rose, E., Cohen, S., & Yahaya, M. A. (1989). Nationwide survey of postsecondary education services for college students with learning disabilities. *Exceptional Children, 56*, 236-245.

Chabot, R. J., Zehr, H. D., Prinzo, O. V., & Petros, T. V. (1984). The speed of word recognition subprocesses and reading achievement in college students. *Reading Research Quarterly, 19,* 147-161.

Educational Testing Service. (1998). *Policy statement for documentation of a learning disability in adolescents and adults* [Brochure]. Princeton, NJ: Office of Disability Policy.

Educational Testing Service. (1986*). Graduate Record Examination*. Princeton, NJ: Educational Testing Service.

Fink, R., Lissner, S., & Rose, B. (1999, July). *Extended Time—When? How Much?* Paper presented at the 22nd Annual Conference of the Association on Higher Education And Disability, Atlanta, GA.

Frauenheim, J. G., & Heckerl, J. R. (1983). A longitudinal study of psychological and achievement test performance in severe dyslexic adults. *Journal of Learning Disabilities, 16,* 339-347.

Fuchs, L., & Fuchs, D. (2001). Helping teachers formulate sound test accommodation decisions for students with learning disabilities. *Learning Disabilities Practice, 16,* 174-181.

Geary, D. C., & Brown, S. C. (1990). Cognitive addition: strategy choice and speed-of-processing differences in gifted, normal, and mathematically disabled children. *Developmental Psychology, 27,* 398-406.

Halla, J. W. (1988). A psychological study of psychometric differences in Graduate Record Examinations General Test scores between learning disabled and non-learning disabled adults (Doctoral dissertation, Texas Tech University, 1988). *Dissertation Abstracts International, 49,* 194.

Hayes, F. B., Hynd, G. W., & Wisenbaker, J. (1986). Learning disabled and normal college students' performance on reaction time and speeded classification tasks. *Journal of Educational Psychology, 78,* 39-43.

Hill, G. A. (1984). Learning disabled college students: The assessment of academic aptitude. (Doctoral dissertation, Texas Tech University, 1984). *Dissertation Abstracts International, 46,* 147.

Hughes, C. A., & Smith, J. O. (1990). Cognitive and academic performance of college students with learning disabilities: A synthesis of the literature. *Learning Disability Quarterly, 13,* 66-79.

Jarvis, K. A. (1996). Leveling the playing field: *A comparison of scores of college students with and without learning disabilities on classroom tests* (Doctoral dissertation, The Florida State University, 1996). *Dissertation Abstracts International*, *57,* 111.

Longstreth, L. E., Walsh, D. A., Alcorn, M. B., Szeszulski, P. A., & Manis, F. R. (1986). Backward Masking, IQ, SAT and reaction time: Interrelationships and theory. *Personality and Individual Differences, 7,* 643-651.

Mellard, D. F. (1990). The eligibility process: Identifying students with learning disabilities in California's community colleges. *Learning Disabilities Focus, 5, 75-90.*

Munger, G. F., & Loyd, B. H. (1991). Effect of speededness on test performance of handicapped and nonhandicapped examines. *Journal of Educational Psychology, 85,* 53-57.

National Board of Medical Examiners. (2001, September). *United States medical licensing examination test accommodations: General guidelines for all test takers with disabilities.* [on-line document]: Hostname: www.nbme.org.

Nelson, R., & Lignugaris-Kraft, B. (1989). Postsecondary education for students with learning disabilities. *Exceptional Children, 56*, 246-265.

Ofiesh, N. S. (1997). Using processing speed tests to predict the benefit of extended test time for university students with learning disabilities. (Doctoral dissertation, The Pennsylvania State University, 1997). *Dissertation Abstracts International, 58,* 76.

Ofiesh, N. S. (1999, July). *Extended test time: A review of the literature.* Paper presented at the 22nd Annual Conference for the Association on Higher Education And Disability, Atlanta, Georgia.

Ofiesh, N. S. (2000). Using processing speed tests to predict the benefit of extended test time for university students with learning disabilities. *Journal of Postsecondary Education and Disability, 14*, 39-56.

Ofiesh, N., Brinkerhoff, L., & Banerjee, M. (2001, July). *Test accommodations for persons with disabilities: No easy answers.* Paper presented at the 24th Annual Conference for the Association on Higher Education And Disability, Portland, OR.

Ofiesh, N., Hughes, C., & Scott, S. (2002). *Extended test time and postsecondary students with learning disabilities: Using assessment data in the decision-making process.* Manuscript in preparation.

Ofiesh, N. S., Kroeger, S., & Funckes, C. (2002, July). *Use of processing speed and academic fluency test scores to predict the benefit of extended time for university students with learning disabilities*. Paper presented at the 25th Annual Conference for the Association on Higher Education And Disability, Washington, DC.

Ofiesh, N. S. & McAfee, J. K. (2000). Evaluation practices for college students with LD. *Journal of Learning Disabilities, 33,* 14-25.

Ragosta, M.. & Wendler, C. (1992). *Eligibility issues and comparable time limits for disabled and nondisabled SAT examinees* (Rep. No. 92-5). New York: College Entrance Examination Board.

Rehabilitation Act of 1973, 29 U.S.C. § 794 et seq.

Runyan, M. K. (1991a). The effect of extra time on reading comprehension scores for university students with and without learning disabilities. *Journal of Learning Disabilities, 24,* 104-108.

Runyan, M. K. (1991b). Reading comprehension performance of learning disabled and non learning disabled college and university students under timed and untimed conditions (Doctoral dissertation, University of California, Berkeley, 1991). *Dissertation Abstracts International, 52,* 118.

Salvia, J., & Ysseldyke, J. (2001). *Assessment* (8th ed.). Boston: Houghton Mifflin.

Use research, develop database to determine extended time policy. (2000, February). *Disability Compliance in Higher Education, 5*(7), p. 8.

Weaver, S. M. (1993). The validity of the use of extended and untimed testing for postsecondary students with learning disabilities (Doctoral dissertation, University of Toronto, Toronto, Canada, 1993). *Dissertation Abstracts International, 55,* 183.

Weaver, S. M. (2000). The efficacy of extended time on tests for postsecondary students with learning disabilities. *Learning Disabilities: A Multidisciplinary Journal*, *10*, 47-55.

Wainer, H., & Braun, H. I. (Eds.). (1988). *Test validity.* Hillsdale, Lawrence Earlbaum.

Wolff, P. H., Michel, G. F., Ovrut, M., & Drake, C. (1990). Rate and timing precision of motor coordination in developmental dyslexia. *Developmental Psychology, 26*, 349-359.

Yost, D. S., Shaw, S. F., Cullen, J. P., & Bigaj, S. J. (1994). Practices and attitudes of postsecondary LD service providers in North America. *Journal of Learning Disabilities, 27,* 631-640.

Ziomek, R. L., & Andrews, K. M. (1996). *Predicting the college grade point averages of special-tested students from their ACT assessment scores and high school grades* (ACT Research Report Series 96-7). Iowa City, IA: American College Testing.

Zurcher, R., & Bryant, D. P. (2001). The validity and comparability of entrance examination scores after accommodations are made for students with LD. *Journal of Learning Disabilities, 34,* 462-471.

About the Authors

Nicole Ofiesh received her Ph.D. in the department of Special Education at Penn State University. She is Assistant Professor of Special Education at the University of Arizona. She has been a special education teacher, learning disabilities specialist, and consultant. Her research interests include the use and validation of assessment tools to determine test accommodations, specifically extended test time, as well as the combined use of learning strategies and technology for secondary and postsecondary students with learning disabilities. She serves as a consultant to the Educational Testing Service (ETS).

Charles Hughes is Professor of Special Education at Penn State University. He has been a general education teacher, special education teacher, state-level consultant, diagnostician, and mainstreaming consultant. His research interests include the development and validation of self-management and learning strategies for adolescents with learning disabilities. He has served on the editorial review board of several learning disabilities journals, was co-editor of JPED, serves as a consultant to ETS and is currently President of CEC's Division for Learning Disabilities.