Learning Disabilities: A Contemporary Journal 3(2), 62–73, 2005

Adult Learning Disabilities Screening Using an Internet-Administered Instrument

Sean Lancaster
Grand Valley State University
Daryl Mellard[1]
University of Kansas

Identifying individuals with specific learning disabilities (SLD) is a complex task, particularly for adult populations. Adult agencies such as vocational rehabilitative services or adult basic education often use different SLD definitions and criteria, are often understaffed, have limited resources, and have a shortage of staff trained on SLD screening (Skinner, Gillespie, & Balkam, 1997). The Adult Learning Disabilities Screening (ALDS) battery (Mellard, 1999) is one of the few valid screening measures for adults currently in use. This study validated an Internet-administered version of ALDS, e-ALDS, by comparing 122 adult education participants who completed the ALDS in paper-and-pencil vs. Internet formats.

Key Words: Learning Disability Assessment, Adult Learning Disabilities, Screening, Computerized Testing, Internet-Based Testing

Adults with undiagnosed specific learning disabiliities (SLD) often experience under- or unemployment, lack of independent living, low self-esteem, and emotional health problems. Thus, high societal and personal costs of failing to recognize and intervene with adults with SLD have been identified in numerous studies (Blackorby & Wagner, 1996; Murray, Goldstein, Nourse, & Edgar, 2000; Sitlington & Frank, 1993; Zigmond & Thornton, 1985). With appropriate intervention, many of these negative outcomes could be overcome. However, in order for these individuals to have access to available support systems, they must first be diagnosed.

Adults with SLD often go undiagnosed because formal assessment opportunities are limited, and even when they are available, the resources to implement assessment are often in short supply. Adult basic education (ABE), vocational rehabilitation (VR), and other social service programs present logical opportunities to assess adults for SLD because they see a large number of people who face multiple barriers to improving their quality of life such as SLD. However, staff at these programs typically do not have training on SLD characteristics or screening methods (Skinner, Gillespie, & Balkam, 1997) and work under rigid time constraints.

Therefore, an efficient and effective approach to screening adults for SLD is needed. A potential solution is computer-based testing, which could reduce the time and expertise needed to administer and score screening batteries, and could be administered anywhere a computer with Internet access is available.

1. Please address correspondence to Daryl Mellard, University of Kansas, Center for Research on Learning, Lawrence, KS 66045. Email: dmellard@KU.edu

Computer-Based Testing

The use of computer-based testing (CBT), whereby computers both present the items and collect responses, has increased steadily since the concept was introduced in the 1950s (Butcher, Perry, & Hahn, 2004). CBT is different from paper-and-pencil tests in that it gives developers the ability to add multimedia, including video and audio, to enhance test items (Schoech, 2001; Stanton, 1999). The increase in the use of CBT has led to research into the equivalence of paper-and-pencil tests and computerized tests. Researchers have generally found the two formats to be equivalent (Bunderson, Inouye, & Greenberg, 1990; Donovan, Drasgow, & Probst, 2000; Neuman & Baydoun, 1998; Vispoel, 2000; Vispoel, Boo, & Bleiler, 2001).

The advantages of using CBT include a reduction in testing time compared to a paper-and-pencil format (Vansickle & Kapes, 1993); elimination of paper while still providing the option to print paper copies if necessary (Cardwell, 2000); more reliability than paper-and-pencil formats (Kapes & Vansickle, 1992; Vansickle & Kapes, 1993); the availability of instant results (Bugbee, 1996; Bugbee & Bernt, 1990; Cardwell, 2000); and participant preference for computerized test formats (Cardwell, 2000; Vispoel, 2000; Zandvliet & Farragher 1997).

However, important disadvantages also exist. One major disadvantage is access to the required technology (Cardwell, 2000). Other disadvantages include the limited computer screen capacity (i.e., small size) compared to paper (Booth, 1998) and the inability to write on test protocols (Kobrin, 2000). Further, Wise and Plake (1990) identified many features of CBT that may influence scores, including the inability to (a) skip items and return to previous items, (b) change previous answers, and (c) review previous items. One of the largest criticisms of CBT is that these three limitations affect an examinee's ability to navigate quickly through a test and change answers when needed (Green, 1988). Thus, Mason, Patry, and Berstein (2001) recommend making computer-based tests so that they more closely mimic paper-and-pencil tests by allowing examinees to go back through the test and change responses.

Remote administration, scoring, and interpretation of results can now occur via computer (Sampson, 2000). According to Jones (1998), the use of Internet-administered testing offers any time-any place access to testing, immediate scoring, and a more limited need for test administrators, which collectively leads to convenient, cost-effective, and efficient testing. Although 50 or more years of research has been conducted on some aspects of computer-based testing, few studies have addressed the Internet as the mode of test delivery.

Technology has long been used to help students with disabilities gain basic skills and knowledge (Boone & Higgins, 1993; Higgins & Boone, 1991; Jones, Torgesen, & Sexton, 1987; MacArthur & Haynes, 1995; Woodward & Gersten, 1992). Yet to date, computer-based assessment of disabilities is not widely practiced and research on its use and effectiveness is scarce in the literature. Few validated, paper-and-pencil SLD screening batteries exist for adults, and at this time no studies have been done to validate a computerized SLD screening battery.

Purpose of the Study

The purpose of this study was to validate an Internet version of ALDS (Mellard, 1999), a previously validated adult SLD screening battery. Specifically, the following

questions were posed regarding comparisons of the paper-and-pencil ALDS, referred to here as p-ALDS, and the computerized, Internet-administered version, called e-ALDS:

1. Do results significantly differ for adults screened via the e-ALDS versus the p-ALDS?
2. Do scores on the e-ALDS for the same participants stay similar over time?
3. Does participants' satisfaction differ with the p-ALDS versus the e-ALDS?

Methods

Research Design

In order to establish the validity of the e-ALDS compared to the p-ALDS, a three-by-two factorial (Group x Session) study was designed. Participants were randomly assigned to one of three groups, each of which participated in two sessions with an interval of approximately one week. Group One completed the p-ALDS in the first session and e-ALDS in the second; Group Two completed the e-ALDS first, followed by the p-ALDS. Finally, Group Three completed the e-ALDS twice. Total scores for the Rating Scale and Inventory were treated as the dependent measures because only these two parts of p-ALDS were included in e-ALDS.

A one-week interval was deemed appropriate for this study's setting—that is, adult education classes that characteristically have short durations and uneven attendance (Moore & Stavrianos, 1995)—in order to obtain repeated measures on all participants. We were concerned that a longer interval would have led to significant attrition, and thus a mortality bias that would threaten the study's validity. Furthermore, we reasoned that since the screening battery is a personal inventory, test memory issues were not relevant.

Participants

Staff from six adult education centers in two midwestern states recruited students to participate. The study population of 122 adult learners was 40% male and 60% female; 77% white, 16% black, 3% Asian, 3% Native American, and 1% Native Hawaiian or Pacific Islander. Age ranged from 18 to 55 years old ($M = 27.8$; $SD = 10.2$).

Each participant signed a consent form signifying agreement to participate in the study. To ensure confidentiality, we do not specify the adult education centers where the study took place.

Examiners

Examiners were selected from staff who had previously been trained on administering ALDS. Each examiner was provided with an instruction manual and a 15-minute training video so that all received the same information on how to conduct the study tasks. The video further explained the manual's key points and demonstrated how to administer e-ALDS. Examiners were encouraged to contact the researchers if any additional questions or concerns developed.

Instruments

Paper-and-pencil ALDS. The Adult Learning Disabilities Screening (ALDS) (Mellard, 1999) is a four-part battery, consisting of self-report paper-pencil items and an interview. Compared to other SLD screening batteries for adults, ALDS has undergone one of the most extensive validation studies to date (Mellard, 1999;

National Adult Literacy and Learning Disabilities Center, 1999; U.S. Department of Education, 2000).

Part 1-Descriptive Information can be independently completed by the examinee. The items assess demographic descriptors, sensory acuities, and English language proficiency.

Part 2-Rating Scale, which contains 25 test items, is designed to be completed independently by participants, yielding self-perception of their behaviors, preferences, and abilities. Specifically, the item content addresses the domains of self-attribution, spelling, reading, social and organizational skills, efficiency, and sense of direction. Participants choose a response by circling the degree to which they agree or disagree with a statement (e.g., When I write, I have trouble putting my ideas in order) or frequency of a behavior (e.g., I organize my tasks). Responses are recorded, and each response is assigned a point value of 1 to 5. The points for each item are added to other items in the same domain to render a domain score. A total score of less than 338 is the cutoff for recommending further diagnostic testing.

Part 3-Inventory, also completed independently, elicits information about health, home, and education. The 19 items address the domains of learning influences, learning problems, educational experiences, mental health, fraction skills, and math operations. Examinees answer either yes or no to questions in each of the six domains (e.g., Have you ever had a head injury?). Scores are tabulated in the same way as the Rating Scale. A score greater than 309 indicates a positive result. Positive scores on either the Rating Scale or the Inventory indicate the need for further testing.

Part 4-Interview is completed with an examiner, who asks questions based on the responses to selected Inventory items. Eight Interview items provide opportunities for clarifying specific Inventory responses. Most examinees complete each part of the battery in less than 8 minutes (Mellard, 1999).

Internet-administered ALDS. The e-ALDS contains only the Rating Scale and Inventory, parts 2 and 3 of ALDS, along with a short tutorial that examinees are required to complete prior to taking the e-ALDS. The e-ALDS runs on any computer with the Macromedia Flash player (Macromedia, 2004), which is automatically installed along with the most common Internet browsers (e.g., AOL, Netscape, MS Internet Explorer) and operating systems (e.g., Apple and Windows XP). The Internet-administered battery presents one item at a time (see sample screen shot in Figure 1) in the same order as p-ALDS. Examinees are free to return to previous items to change answers. Audio clips of each item and answer choices are available. (Examiners are able to determine afterward the extent to which this audio feature was used.)

Satisfaction survey. A satisfaction survey in paper-and-pencil format was designed to determine whether participants were satisfied with the p-ALDS and e-ALDS and their preference between the two versions. The survey was completed for all groups and sessions.

The survey contained the following items: (a) I am satisfied with the format of the battery of questions I just completed; (b) The directions were clear and easy to understand; (c) Each item was clear and easy to understand; (d) Another person helped me to complete the battery; and (e) The amount of time the battery took was about right. Participants rated each item on a Likert scale with five choices ranging

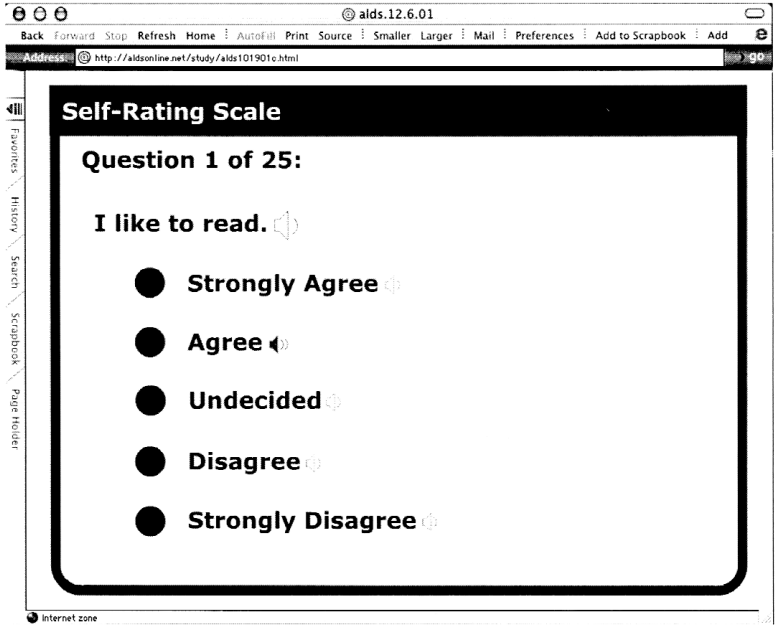

Figure 1. Screen shot: Question 1 of rating scale.

from 1 = Strongly Agree to 5 = Strongly Disagree. In the second session, participants who completed both the p-ALDS and e-ALDS were given an additional item asking them to circle the version of the battery they preferred.

Procedures

Participants were randomly assigned to Group One, Two, or Three in each participating adult education center. The examiners provided the materials, computer and Internet connection for data collection. The first author provided assistance to examiners via phone and email to overcome unanticipated barriers. To cut down on administration time, examiners were encouraged to test small groups of fewer than six participants while ensuring each participant was able to complete the battery in a quiet environment.

Examiners administered the p-ALDS or e-ALDS battery to participants based on group assignments. After a one-week interval, examiners again administered the p-ALDS or e-ALDS battery depending on group assignment. Examiners also administered the satisfaction survey with all participants.

Data Analysis

Multivariate analysis of variance was selected as the statistical test for effects. The data were treated as a 3 (Groups) by 2 (Sessions) factorial design. Sessions served as a repeated measure, allowing each participant to serve as his/her own control. The two dependent variables were the participants' scores on the ALDS Rating Scale and the Inventory. Alpha was set at .05 for testing the hypothesis of mean differences.

Results

Validation Results

The significant validity question was whether one group was advantaged or disadvantaged by their group membership. That is, a statistically significant interaction of Group x Session might cause concern about the accuracy of the scores. Tables 1 and 2 present the results of the multivariate analysis of variance for the Rating Scale and Inventory total scores. The results indicated no reliable difference in the interaction of Group with the Session, Wilks' Lambda = .96, $F(4, 230) = 1.193$, $p = .315$, or the change in scores between the two sessions, Wilks' Lambda = .991, $F(2, 115) = .518$, $p = .597$.

Table 1

Descriptive Statistics from Rating Scale and Inventory

	Mean	SD	Minimum	Maximum
		Rating Scale		
Group One				
Session 1: p-ALDS	347.36	11.137	318	368
Session 2: e-ALDS	347.41	10.472	317	367
Group Two				
Session 1: e-ALDS	349.88	8.648	331	362
Session 2: p-ALDS	350.15	8.310	332	366
Group Three				
Session 1: e-ALDS	353.35	12.653	312	371
Session 2: e-ALDS	353.63	12.790	312	373
		Inventory		
Group One				
Session 1: p-ALDS	300.77	11.672	287	325
Session 2: e-ALDS	302.56	10.765	289	325
Group Two				
Session 1: e-ALDS	301.00	12.504	288	334
Session 2: p-ALDS	300.35	12.186	287	332
Group Three				
Session 1: e-ALDS	294.93	6.833	289	310
Session 2: e-ALDS	294.95	7.084	287	311

Note. Group One, $n = 39$; Group Two, $n = 40$; Group Three, $n = 43$.

Table 2

MANOVA Results: Wilks' Lambda

	Value	F^a	effect *df*	error *df*	*p*	η^2
			Between Subjects			
Intercept	.000	193962	2	115	.000	1.000
Group	.909	2.820	4	230	.026	.047
			Within Subjects			
Session	.991	.518	2	115	.597	.009
Session x Group	.960	1.193	4	230	.315	.020

Note. [a] Exact statistic; *p* = Significance; η^2 = Partial eta squared.

A statistically significant main effect was found for Group, indicating that the group membership variable accounted for some differences in scores, Wilks' Lambda = .909, $F(4, 230) = 2.82$, $p = .026$. An inspection of Table 1 shows that the e-ALDS/e-ALDS group had higher mean values than the other two groups. This difference is attributed to sampling variations that were not resolved through random assignment of participants.

Although the interaction term of Group x Session was not statistically significant, effect sizes were calculated to obtain added information for assessing the magnitude of the observed mean differences. Cohen's *d* formula (Cohen, 1977) was used to evaluate the standardized difference between the three groups' scores for the first and second administration of the ALDS. The calculated *d* values of effect size are reported in Table 3. Cohen's (1977) guidelines for interpreting effect sizes are helpful when considering the results. Cohen considered *d* values of 0.2 as "small" and 0.5 as "medium."

Our calculated *d* values demonstrated a significant range in magnitude, with the largest effect ($d = 0.2659$) for the Inventory score of Group One, which completed the administration order of the ALDS paper version and computer-based version, respectively. The second highest effect size value was for Group Two, which completed the Inventory in the opposite order (i.e., e-ALDS and then p-ALDS). For Group One (p-ALDS and e-ALDS), the mean scores increased, whereas for Group Two (e-ALDS and p-ALDS), the mean scores decreased between the two administrations. These differences could reflect the sample size and sampling variability. In both cases the e-ALDS yielded a higher mean value. Finally, Group Three, which was administered the e-ALDS for both Inventory administrations, showed no reliable mean difference and had the smallest effect size. The e-ALDS version does not change the screening outcome. As found by way of the null hypothesis tests, observed differences do not vary beyond sampling fluctuations.

Table 3
Effect Size Calculations

	Group Comparisons		
	Group One (p-ALDS & e-ALDS)	Group Two (e-ALDS & p-ALDS)	Group Three (e-ALDS & e-ALDS)
Rating Scale			
Cohen's *d*	-0.012125636	-0.068594833	-0.098319101
Inventory			
Cohen's *d*	-0.265978311	0.165967073	-0.004964405

A second validity question was whether the scores significantly changed from the first to the second administration after the one-week interval. Since the main effect of Session was not statistically significant, no reliable change was measured.

A practical question is whether classification results differed between administrations on the p-ALDS Rating Scale and Inventory compared to the e-ALDS Rating Scale and Inventory for *each group,* that is, whether the classification results were stable. For Group One, 35 of 39 participants were classified identically after both administrations, for an agreement of 89.8%. For Group Two, agreement was 92.5% and 97.5% on the Rating Scale and Inventory, respectively. For Group Three, the

Rating Scale agreement was 100% and 95.3% for the Inventory. Thus, for all three groups almost identical referral decisions were made.

Internet-Administered ALDS Reliability

In order to establish consistency of the results on the e-ALDS, temporal reliability was measured by comparing Group Three results from the first administration to those of the second administration on both the Rating Scale and the Inventory. Pearson correlations revealed a high correlation between the Group Three Summed Total Part scores from session one to session two on both the Rating Scale ($N = 43$, $r = .974$) and the Inventory ($N = 43$, $r = .838$).

Satisfaction

Participants were generally satisfied with both versions of the battery and did not appear to require much assistance to complete either version (see Table 4). Between-group comparisons were done on the satisfaction rates reported by the three groups for the e-ALDS only. One-way ANOVAs were run for each item on the satisfaction questionnaire. Results indicated no significant differences between the satisfaction ratings for Item 1, $F(2, 97) = 2.85$, $p = .063$, and Item 5, $F(2, 97) = 1.05$, $p = .355$. However, significant differences were found between the satisfaction for Item 2, $F(1, 97) = 6.32$, $p < .05$, Item 3, $F(1, 97) = 6.28$, $p < .05$ and Item 4, $F(1, 97) = 6.87$, $p < .05$. For Item 2, the Bonferroni Test revealed significant differences at the .05 level between Groups One and Three, indicating that participants in Group Three were more satisfied with the clarity and ease of the e-ALDS directions than were participants in Group One. For Item 3, the Bonferroni Test revealed significant differences at the .05 level between Groups One and Three and Groups Two and Three, suggesting that participants in Group Three were more satisfied with the clarity and ease of the items than were participants in Groups One and Two. Finally, for Item 4, the Bonferroni Test revealed significant differences at the .05 level between for Groups One and Two suggesting that participants in Group Two received more help while completing the assessment using e-ALDS.

Additional within-group comparisons were done with Groups One and Two to determine whether participants in either of these groups were equally satisfied with both versions of the battery. For Group One, *t* tests revealed no significant differences between satisfaction ratings of participants for Items 1 through 5 on either version. For Group Two, *t* test revealed significant differences between satisfaction ratings of

Table 4
Satisfaction Survey Results

	Group One				Group Two				Group Three	
	p-ALDS		e-ALDS		e-ALDS		p-ALDS		e-ALDS	
Item	Mean	*(SD)*	Mean	*(SD)*	Mean	*(SD)*	Mean	*(SD)*	Mean	*(SD)*
1	1.97	(.67)	1.84	(.69)	1.56	(.50)	1.63	(.55)	1.48	(.68)
2	1.53	(.68)	1.42	(.50)	1.19	(.40)	1.26	(.44)	1.06	(.25)
3	1.57	(.63)	1.52	(.51)	1.44	(.50)	1.29	(.46)	1.13	(.34)
4	4.53	(.86)	4.39	(.84)	4.92	(.28)	4.37	(1.09)	4.71	(.53)
5	1.73	(.83)	1.71	(.82)	1.58	(.65)	1.51	(.61)	1.45	(.62)

Note. 1 = Strongly Agree and 5 = Strongly Disagree.

participants for Item 4, indicating that more participants received help while completing the p-ALDS than while completing the e-ALDS; it is worth noting that participants in Group Two completed the e-ALDS first, followed by the p-ALDS.

Finally, after participants in Groups One and Two completed both versions of the battery, they were asked in Item 6 which version they preferred. Of the 31 participants in Group One, 6 preferred the p-ALDS and 25 preferred the e-ALDS. Of the 34 participants in Group Two, 13 preferred the p-ALDS while 21 preferred the e-ALDs. Results of a chi-square demonstrated that significantly more participants preferred the e-ALDs than the p-ALDS, $\chi^2(1, 66) = 11.88$, $p < .001$.

Discussion

The central question in this study was whether an Internet-administered version of the ALDS would yield valid results. First, for all groups, scores on the p-ALDS and e-ALDS were comparable, with no significant differences due to mode or order of administration of the Rating Scale and the Inventory. This finding suggests that practitioners can feel comfortable treating the results of both modes of assessment as equivalent.

The results of the two administrations of the e-ALDS for Group Three showed no significant differences between the results of the first and second administration for both the Rating Scale and the Inventory. Thus, the e-ALDS appears to yield stable scores over time.

Thirty-two percent of Group One and Two participants, all of whom attended adult basic education programs, screened positive for a potential SLD when using p-ALDS. This is similar to the 28.7% who screened positive in the original ALDS validation study (Mellard, 1999). This latter percentage included individuals from sites other than adult basic education sites such as individuals receiving social rehabilitation services and a small number of incarcerated individuals. Additionally, 22% of participants who took the e-ALDS screened positive, closely approximating the results of the original validation study and the number of individuals who screened positive on the p-ALDS in the current study.

In general, participants were satisfied with both versions of the battery. Participants in Groups One and Two were more satisfied with the clarity of the directions and items on the e-ALDS than the p-ALDS. Considerably more participants in Groups One and Two preferred the e-ALDS to the p-ALDS. These results are consistent with other studies indicating a preference for computerized testing over paper-and-pencil testing (Vispoel, 2000; Zandvliet & Farragher, 1997).

Limitations

Despite these positive findings, the study suffers from at least two limitations. First, low computer proficiency or a lack of comfort with the computer may have affected participants' accuracy and, thus, their scores. In practice, an examiner would want to ensure that the client is comfortable with the technology and proficient with the task demands (e.g., changing answers).

Second, the potential benefits of reduced administration and scoring time were not measured, and thus were not compared to the paper-and-pencil version. These data

could have lent greater support to use of the e-ALDS in adult basic education centers or social and rehabilitative services offices that are understaffed and/or where existing staff members are not trained to provide screening for SLD (Skinner et al., 1997).

Conclusion

This research has extended the literature on the use of screening batteries for adults with possible SLD. Current findings demonstrate that an Internet-administered form of the ALDS screening battery, e-ALDS, (a) can be used to screen adults for SLD with results similar to those of the paper-and-pencil version; (b) is reliable over time; and (c) elicits satisfaction ratings equal to or better than those of the paper-and-pencil version.

These findings have important implications for adult-focused social service agencies and education programs such as agencies working with welfare recipients or persons receiving Temporary Assistance to Needy Families (TANF). The ALDS provides a valuable screening battery that places minimal requirements on staff time, provides nearly instant assessment results, and yields information that can help decide whether further SLD assessment is warranted. With resources scarce, the e-ALDS offers an effective and efficient SLD screening method that may also prove useful in postsecondary education settings.

Sean Lancaster *is an Assistant Professor at Grand Valley State University in Grand Rapids, Michigan.* ***Daryl Mellard*** *is a Research Associate at the University of Kansas, Center for Research on Learning in Lawrence, Kansas. With several colleagues he is engaged in research and evaluation activities generally focusing on making adult community services more inviting, welcoming, and inclusive for persons with disabilities.*

Daryl Mellard's work on this article was supported by the USDE/OSEP project Individual Accommodation Model in Postsecondary Settings (H324M980109) and the USDE/NIDRR project Rehabilitation Research and Training Center on Full Participation in Independent Living (H133B000500).

References

Blackorby, J., & Wagner, M. (1996). Longitudinal post-school outcomes of youth with disabilities: Findings from the national longitudinal transition study. *Exceptional Children, 62*(5), 399–413.

Boone, R., & Higgins, K. (1993). Hypermedia basal readers: Three years of school-based research. *Journal of Special Education Technology, 12*(2), 86–106.

Booth, J. (1998). The user interface in computer-based selection and assessment: Applied and theoretical problematics of an evolving technology. *International Journal of Selection and Assessment, 6*(2), 61–81.

Bugbee, A. (1996). The equivalence of paper-and-pencil and computer-based testing. *Journal of Research on Computing in Education, 28*(3), 282–299.

Bugbee, A., & Bernt, F. (1990). Testing by computer: Findings in six years of use 1982–1988. *Journal of Research on Computing in Education, 23*(1), 87–100.

Bunderson, C., Inouye, D., & Greenberg, A. (1990). *Computers in educational assessment: An opportunity to restructure educational practices.* A paper prepared for the U.S. Congressional Office of Technology Assessment. (ERIC Document Reproduction Service No. ED340771)

Butcher, J., Perry, J., & Hahn, J. (2004). Computers in clinical assessment: Historical developments, present status, and future challenges. *Journal of Clinical Psychology, 60*(3), 331–345.

Cardwell, K. (2000). Electronic Assessment. *Learning and Leading with Technology, 27*(7), 22–26.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* New York: Academic Press, Inc.

Donovan, M., Drasgow, F., & Probst, T. (2000). Research reports–Does computerizing paper-and-pencil job attitude scales make a difference? New IRT analyses offer insight. *Journal of Applied Psychology, 85*(2), 305–313.

Green, B. (1988). Critical problems in computer-based psychological measurement. *Applied Measurement in Education, 1*(3), 223–241.

Higgins, K., & Boone, R. (1991). Hypermedia CAI: A supplement to an elementary school basal reader program. *Journal of Special Education Technology, 11*(1), 1–15.

Jones, M. (1998). *Creating electronic learning environments: Games, flow, and the user interface.* Urbana-Champaign: University of Illinois. (ERIC Document Reproduction Service No. ED423842)

Jones, K., Torgesen, J., & Sexton, M. (1987). Using computer guided practice to increase decoding fluency in learning disabled children: A study using the Hint and Hunt 1 program. *Journal of Learning Disabilities 20,* 122–128.

Kapes, J., & Vansickle, T. (1992). Comparing paper-pencil and computer-based versions of the Harrington-O'Shea career decision-making system. *Measurement and Evaluation in Counseling and Development, 25*(1), 5–13.

Kobrin, J. (2000). *An investigation of the cognitive and paper-and-pencil reading comprehension test items.* (ERIC Document Reproduction Service No. ED442836)

MacArthur, C., & Haynes, J. (1995). Student assistant for learning from text (SALT): A hypermedia reading aid. *Journal of Learning Disabilities, 28*(3), 150–159.

Macromedia. (2004). *Macromedia Flashplayer Statistics.* Retrieved December 14, 2004, from http://www.macromedia.com/software/player_census/flashplayer/

Mason, B., Patry, M., & Berstein, J. (2001). An examination of the equivalence between non-adaptive computer-based and traditional testing. *Journal of Educational Computing Research, 24*(1), 29–39.

Mellard, D. (1999). *The Adult Learning Disabilities Screening.* Perry, KS: Kaw Valley Resources, Inc.

Moore, M., & Stavrianos, M. (1995). *Review of adult education programs and their effectiveness: A background paper for reauthorization of the Adult Education Act.* Washington, DC: Mathematica Policy Research, Inc.

Murray, C., Goldstein, D., Nourse, S., & Edgar, E. (2000). The postsecondary school attendance and completion rates of high school graduates with learning disabilities. *Learning Disability Research & Practice, 15*(3), 119–127.

National Adult Literacy and Learning Disabilities Center. (1999). *Bridges to practice: A research-based guide for literacy practitioners serving adults with learning disabilities, Guidebook 2, The assessment process.* Washington, DC: Author.

Neuman, G., & Baydoun, R. (1998). Computerization of paper-and-pencil tests: When are they equivalent? *Applied Psychological Measurement, 22*(1), 71–83.

Sampson, J., Jr. (2000). Using the Internet to enhance testing in counseling. *Journal of Counseling & Development, 78*(3), 348–56.

Schoech, D. (2001). Using video clips as test questions: The development and use of multimedia exam. *Journal of Technology in Human Services, 18*(3-4), 117–131.

Sitlington, P., & Frank, A. (1993). Dropouts with learning disabilities: What happens to them as young adults? *Learning Disabilities Research and Practice, 8*(4), 244–252.

Skinner, L., Gillespie, P., & Balkam, L. (1997). Adults who learn differently: Help through a volunteer literacy program. *Annals of Dyslexia, 47,* 185–202.

Stanton, J. (1999). Validity and related issues in web-based hiring. *The Industrial Psychologist, 36*(3), 69–77.

U.S. Department of Education. (2000, April). *Learning disabilities and Spanish-speaking adult populations: The beginning of a process.* Paper presented at the meeting of Division of Adult Education and Literacy, Office of Vocational and Adult Education, U.S. Department of Education, San Antonio, TX.

Vansickle, T. R., & Kapes, J. T. (1993). Comparing paper-pencil and computer-based versions of the Strong-Campbell Interest Inventory. *Computers in Human Behavior, 9*(4), 441–449.

Vispoel, W. (2000). Computerized versus paper-and-pencil assessment of self-concept: Score comparability and respondent preferences. *Measurement and Evaluation in Counseling and Development, 33*(3), 130–143.

Vispoel, W., Boo, J., & Bleiler, T. (2001). Computer and paper-and-pencil versions of the Rosenberg Self-Esteem Scale: A comparison of psychometric features and respondent preferences. *Educational and Psychological Measurement, 61*(3), 461–474.

Wise, S., & Plake, B. (1990). Computer-based testing in higher education. *Measurement and Evaluation in Counseling and Development, 23*(1), 3–10.

Woodward, J., & Gersten, R. (1992). Innovative technology for secondary students with learning disabilities. *Exceptional Children, 58*(5), 407–421.

Zandvliet, D., & Farragher, P. (1997). A comparison of computer-administered and written tests. *Journal of Research on Computing in Education, 29,* 423–438.

Zigmond, N., & Thornton, H. (1985). Follow-up of post-secondary age learning disabled graduates and dropouts. *Learning Disabilities Research, 1,* 50–55.

Received September 26, 2004
Revised January 10, 2005
Accepted January 21, 2005