USING THE INTERNET FOR CAREER-TECHNICAL ASSESSMENT: A PILOT PROJECT IN OHIO

James T. Austin
Robert A. Mahlman
Ohio State University

ABSTRACT

Occupation-specific assessment over the Internet was evaluated in a pilot project that addressed four objectives. One hundred nineteen students in seven classes/schools took an administrative office technology assessment given on demand over the Internet. Sixty-one students had taken the same assessment in paper form six to eight weeks earlier. The scores for Internet administration of that subgroup were significantly higher (4.85) but exhibited similar variability and internal consistency reliability. Responses to a background and reaction survey indicated few problems with testing (accessing, downloading, taking, and submitting the test, plus understanding feedback). A majority of students responded positively to global evaluations (75% reported a positive experience; 62% preferred or greatly preferred Internet format over the traditional). Results suggest that the advantages of testing via the Internet outweigh the disadvantages. Strategies and tactics for moving vocational assessment to the Internet are discussed.

It is increasingly mandated that educators demonstrate the results of instruction, given the proliferation of standards (content and performance) in education (Resnick & Wirt, 1996). The requirement to demonstrate results is especially important in career and technical education given substantial federal, state, and local investments. The Ohio Vocational Competency Assessment (OVCA) is a competency-based, content valid system that helps to meet this requirement.

BACKGROUND

The OVCA system measures specific occupational and general employability knowledge that result from career-technical instruction and related experiences. The OVCA system consists of three components and is administered annually with results reported to students, teachers, and administrators. The first of three components is 39 occupation-specific assessments. The second and third OVCA components are both general. One is a cross-occupational assessment, Employability Skills, and the other is a battery consisting of three of four ACT (formerly American College Testing) WorkKeys assessments. The specific assessments vary with occupational taxonomies. For example, program completers in the Administrative Office Technology curriculum, which is part of the Business cluster, were the focus of this project. Individual learners whose school districts elect to use the OVCA system might take as many as five assessments during an

annual testing window. The particular instruments used would be the occupation-specific Administrative Office Technology and a cross-occupational Employability Skills (CETE) assessment, plus three WorkKeys (Applied Mathematics, Reading for Information, and Locating Information). While this traditional assessment system exhibits good psychometric quality in terms of reliability, content validity, and reasonable benchmark scores, there are potential advantages to a secure computer system capable of operating at a distance. One way to realize these advantages is through strategic use of the Internet.

Hambleton (1989, 1996), Drasgow, Olson, Keenan, Moberg, and Mead (1993) and Sands, Waters, and McBride (1997) have advocated the use of computers in educational and organizational assessment. Sands et al. (1997), for example, edited a volume that explores the evolution of the Armed Services Vocational Aptitude Battery from a traditional, paper-pencil assessment to one that is largely administered by computer and uses adaptive programming to present items to test takers. Drasgow and Olson-Buchanan (1999) edited a book that examines innovations in computerized assessment across multiple domains (e.g., personnel selection, scholastic assessment, and certification). Kingsbury and Houser (1999) presented their development and implementation of computerized assessment for the Portland City Schools. Zakrzewski and Bull (1998) presented a description of their planning and implementation (1994-1997) of summative and formative assessments at a university in England.

INTERNET IMPORTANCE AND EXPANSION

Members of the educational community are aware of the fundamental changes in communication and commerce that are occurring as a result of the explosive expansion of the Internet. Physically, the Internet and World Wide Web (WWW) are a connected network of computers that enable communication and commerce at a distance (Hahn & Stout, 1994; Maran, 1995; Thomas, 1995). It is conceptually, however, that the influence of the Internet is vast (Feldman & Krumenaker, 1995; Institute for Information Studies, 1997). The Internet creates a fundamental shift in how the enterprise of education is conducted. Curriculum, administration, and assessment are all being influenced and that influence is growing.

TESTING OVER THE INTERNET

In 1997, the Pennsylvania Department of Education sponsored a pilot study to evaluate use of the Internet for test delivery. The Pennsylvania project is reported in detail in a technical report (Slivinski, Hardwicke, Kapes, Boyer, Ip, & Martinez, 1997) and in two articles (Bicanich, Hardwicke, Slivinski, & Kapes, 1997; Kapes, Martinez, Ip, Slivinski, & Hardwicke, 1998). The articles presented evaluation of several research questions. In brief, the Pennsylvania project involved two research phases and included 14 schools and over 370 students. The Vocational Technical Educational Consortium of States (VTECS), a consortium that provides members with competency-based, career-technical outcome standards, curriculum resources, and assessment vehicles, was the source of the test items. The computer served as a "page-turner," which means that the computer presents and scores the test, but does not adapt the items to examinees based on their responses as in computer adaptive testing (CAT). The first phase consisted of students from Computer Repair and Computer Specialist curricula (N=160), who completed randomly generated sets of VTECS test items for self-assessment and instructional purposes. The second phase involved testing students in Child Care and Auto Body Repair vocational programs (N=360) using both paper-pencil and Internet formats for comparison purposes. The design was a test-retest equivalency with control groups. In addition to the test performance comparisons for the second

design, short attitude surveys were completed by students, teachers, and test administrators. The results indicated, first, no statistically significant differences in test performance across groups defined by demographic and special needs status. Second, the Internet provided a cost-effective alternative to the traditional assessment format. Third, students preferred Internet testing by a 3-1 margin. Fourth, a qualitative evaluation of technical feasibility indicated no insoluble implementation problems. Finally, test security was not adversely affected by the Internet format, although this conclusion must be qualified by the fact that the VTECS item bank itself is not totally secure.

PURPOSE AND OBJECTIVES

Two potential motivations to develop and evaluate testing systems are *responsiveness to stakeholders* (speed of scoring, printing/mailing cost savings) and *responsiveness to advancing technology* (use of computers, the Internet, and evolution toward computer adaptive testing). There are a number of advantages to assessing students in this manner. Because assessment specialists are interested in methods that provide additional value to stakeholders (e.g., more rapid feedback to test takers, cost reduction, and security enhancement), the present project investigated the feasibility of testing over the Internet. The primary research objectives were:

1. To examine the descriptive and psychometric characteristics of the test delivered via the Internet (mid-May),
2. To compare performance on the Internet format test to a paper-pencil version of the test that the students had taken (in March),
3. To assess student attitudes about Internet testing, and
4. To evaluate technical feasibility.

LIMITATIONS

Given our smaller sample and choice of program, we were unable to examine the effects of various demographics on test score differences. We were also unable to have sufficient experimental control to permit counterbalancing of administration.

METHODS

We collaborated with a consulting organization on a pilot evaluation of use of the Internet to deliver vocational testing. The project was funded by the state's Department of Education and was conducted during April-May 1998. The project involved testing students from multiple schools using the standard paper-pencil Administrative Office Technology assessment converted to Internet delivery and administering a survey measuring background and reaction variables.

PROJECT DESIGN

The Ohio project differed from the Pennsylvania project in some ways, such as the use of different test items. Also, the Ohio project was smaller than the Pennsylvania project in terms of schools and participants. However, the crucial similarity was evaluation of use of the Internet to deliver testing. The objectives, as previously expressed, were similar. The consulting firm involved in the project produces Internet computer programming for educational usage, including assessment and curriculum applications. The Pennsylvania testing project used software written in PERL while the present project, due to security concerns, was written and implemented using a JAVA applet approach. JAVA is a programming language and applets are small programs that can be used as building blocks.

INTERNET TESTING PROCEDURES

The general sequence of Internet testing is as follows. First, examinees access the Website operated by the provider and choose whether to take a tutorial (which could have been taken anytime prior to testing). When the student indicates her or his readiness, the complete test is downloaded from the server. The individual moves through the test responding to the items and is permitted to review responses. When ready, the test is submitted to the central server for scoring and return of feedback to the test-taker. For this study, the computer was considered a "page-turner" to administer and score the assessment. This usage corresponds to what Hambleton (1996) termed 'the earlier use of computers in assessment', but we assert that the benefits still outweigh the costs.

SECURITY ISSUES AND SOLUTIONS

Security is a crucial issue for the OVCA system, given the investment in its development and its high-stakes nature. We suggest that the best approach to test security is to use multiple methods of verifying identities and monitoring test-takers. Obviously, passwords for test takers are a fundamental attribute of computerized test security. We used a database program to create randomized passwords that conveyed information about schools, teachers, and students in 3 sets of 4-digits. One useful feature of the system is that it can be set up so that each password is only valid one time. An anecdote illustrates the utility of this feature. At one of the pilot schools, a teacher distributed the passwords in advance of the testing session because he believed (incorrectly) that a password was required to access the on-line tutorial. On the day of testing, three students reported that their password did not work, which suggests possible attempts to access the testing site. During larger-scale usage of the system, it will be possible to record the location from which the Website was accessed and evaluate such hypotheses.

A second security concern for a distributed testing system pertains to the capability of test takers or others at a remote site to either print or save test materials from their computer stations. Proctors can help to eliminate some such problems, but there is a cost tradeoff associated with increased personnel. JAVA applets enabled the avoidance of these problems by temporarily disabling the print and save features of the browsers (e.g., Netscape) being used during testing sessions. Also, it is possible to store the location (Internet Portal address) from which an individual accesses the system in order to provide an "electronic trail." Finally, a third method of enhancing test security was the use of "windows" during which the server would permit access from specified schools. Thus, schools signed up for primary and alternate windows. Thus, security was approached in this project using multiple methods to trade off strengths and weaknesses.

ADMINISTRATIVE OFFICE TECHNOLOGY (AOT) ASSESSMENT

The instrument used was identical to the paper-pencil version that had been administered as part of the standard testing window during March-April 1998. The same form was used to provide a comparison across formats, although the design control to make the same comparisons made in the Pennsylvania project was absent. The 100-minute assessment contained 100 four-option multiple choice items, including seven items that involved graphics (e.g., tables or charts). Descriptive and psychometric features of the paper-pencil assessment from 1995-1998 are presented in the results section.

TRAINING FOR TEACHERS AND TEST ADMINISTRATORS

Approximately one week before the scheduled testing window of May 7-15, 1998, we conducted a training session for participating schools at a local Joint Vocational School. The authors, personnel

from the state Department of Education, and the consultant were present. At this session, held in a computer lab, we presented the rationale for Internet testing, a general overview of the system (including security issues), and details of the specific project. All attendees used a computer to access the Website and worked through an early version of the tutorial and a practice test. Several suggestions were made at the session and implemented by the consultant. Following training, several individuals stated that they believed the process could be extremely helpful in improving assessment by speeding up the feedback process and permitting assessment later in the school year.

SCHOOL AND STUDENT SAMPLE

Eleven schools (e.g., Joint Vocational Schools, Career Centers) indicated willingness to participate at the start of the project. Due to various problems, mostly technological in nature (e.g., software firewalls prohibiting access), the final set consisted of seven schools. The participating institutions represented a geographic cross-section of the state. Both urban and rural schools were present in the sample, although urban/suburban schools predominated. The student sample consisted of 141 students who participated in the study and were assigned passwords. One hundred nineteen students took the computerized assessment (performance data). However, only about 100 students completed the background and reactions survey (attitudinal data), perhaps because of its positioning at the end of the assessment.

DATA ANALYSIS

Several analyses were conducted to address the objectives of the study. Various statistical aspects were examined, including (a) item/test statistics for both paper-pencil and Internet versions, (b) responses to the attitudinal surveys by test takers, and (c) relationships of these variables with demographics. Descriptive and psychometric statistics were calculated for usable responses to the 100-item AOT assessment and the 16-item attitude survey. Correlational analyses were performed to examine the associations among test scores (paper-pencil and Internet), attitudinal responses, and demographic variables. The objective pertaining to feasibility, specifically the logistical and technical issues surrounding implementation, was addressed through telephone interviews with teachers.

RESULTS

Table 1 contains results for the paper-pencil assessment for 1995-1998, based on samples of 3,000 test-takers each year (3,500 take the test each year). The forms for 1997 and 1998 were kept identical to facilitate reporting and the establishment of benchmark scores.

Table 1

Descriptive and Psychometric Features of the Paper-Pencil Assessment, 1995-1998

Year	Mean	Range	Std Deviation	Cronbach α
1995	48.60	0-87	13.63	.886
1996	54.01	1-91	15.46	.927
1997	55.85	0-93	17.22	.924
1998	58.28	1-93	16.82	.920

Descriptive data related to the first research objective are provided in Table 2 and the subsequent paragraph. The demographic distribution of this sample indicates a majority of females (92.4%)

and Caucasians (88.1%). All but two persons reported English as their first language. Approximately two-thirds of the students reported being in a two-year vocational curriculum (67.2%). Specific sample characteristics in Table 2 were comparable to the demographic profile of 3,436 individuals taking the 1998 paper-pencil assessment. As is typical with field data, there were missing responses for test items and for reaction survey questions. Given the small sample size, all available data were used for the analyses.

Table 2

Characteristics of Students Taking the Internet Assessment

Variable	Levels	Present Sample	1998 Program Population
Gender			
	Male	9(07.6%)	402 (11.3%)
	Female	110 (92.4%)	3137 (88.2%)
Ethnicity			
	African-American	11 (09.2%)	440 (12.4%)
	Euro-American	105 (88.1%)	2497 (70.2%)
	Other	2 (01.7%)	54 (01.5%)
	Multiracial	1 (00.8%)	59 (01.7%)
Vocational Program			
	1st of 2-year	10 (08.4%)	304 (08.5%)
	2nd of 2-year	70 (58.8%)	1780 (50.0%)
	1st of 1-year	36 (30.3%)	1187 (33.4%)
	Adult	3 (02.5%)	184 (02.5%)

Due to space limitations, the following data with respect to the assessment itself is presented in narrative form. A majority (91%) of the sample attempted 90 or more items (which reduced the variance of this variable). The number of right scores on the computer-administered version fell between 15 and 91 (N=119). The average score was 65.42 with a standard deviation of 13.63. The cut points for dividing the score distribution into quartiles were 58.3, 67.5, and 75.0. The estimated Cronbach's α was .90. These values are comparable with those presented for 1995-1998 in Table 1. The one exception was the mean score on the assessment, which exhibited an increase of 4.4 points for the Internet group compared to their previous paper-pencil scores. This finding was not unexpected given that the Internet assessment was a retest in a different format at an interval of two months.

To address the second research objective, scores on the paper-pencil and computer formats were evaluated for a subset of the students taking the Internet assessment. Not all students take the paper-pencil assessment because the decision to test is made by school districts. We identified 61 students out of 119 (50.4%) who had item scores for both paper-pencil and Internet assessments. For this subset, we calculated means, standard deviations, and Cronbach α internal consistency estimates for both administration formats. The results of this comparison are presented in Table 3 and additional data related to the second research objective is presented in narrative form in the subsequent paragraph.

Table 3

Comparison of the Internet and Paper-Pencil Formats (N=61)

Paper-Pencil	Internet
59.33	64.18
14.81	13.93
2.3	15
82	85
.91	.90

Again due to space limitations, the following data with respect to comparing performance on the Internet format test to a paper-pencil version of the test is presented in narrative form. Another way to view the Internet version is as a retest with an interval of nearly two months, the correlation between the two scores for the subsample was .60. A paired t-test indicated that the difference in means was statistically significant (t = -2.94, df = 60, p = .005). A second t-test was conducted to compare the Internet number of right scores of two groups, those who had taken (N=61) with those who had not taken (N=60) the earlier paper-pencil assessment. The results of that test were not significant (t=1.20, df=119, p=.23). The mean difference of 2.95 favored those without prior experience but was not statistically significant.

ONLINE SURVEY RESULTS

An associated on-line survey (see Appendix) addressed the third research objective. The survey, answered in part by 101 individuals taking the test (complete data were available for 65), was informative about the sample and their reactions to the Internet testing process. The survey questions used either Yes-No, 3-point, or 5-point anchored responses to investigate background and experience, the Internet testing process, and overall reactions. Nearly a third of the sample indicated that they had received some assistance from a facilitator. Four questions pertained to computer access and experience (e.g., computer at home, daily access at school, self-evaluation of prior computer experience, and prior Internet experience). Approximately half of the sample reported that they accessed a computer at home, while 95% reported having daily access at school. Again, most individuals (87 or 95%) reported at least moderate experience with the computer, but far fewer (16%) reported moderate or great experience with the Internet.

With respect to the experiences reported by students on the day of testing, survey questions asked about accessing the site, using the mouse to answer items or navigate through the test, reading the items on the screen, submitting the answers for scoring, receiving and printing the results, having keyboarding problems, understanding the computer-provided results, and experiencing computer anxiety during testing. Only 17 of 92 students (16%) reported major or minor problems accessing the website or downloading the test, whereas 95% reported no problems with using the mouse to answer items or navigate through the screens. Twenty-one of 91 students (23%) reported minor or major problems reading the items on the computer screens. This pattern was repeated for the items pertaining to submitting answers for scoring, receiving answers, and understanding the results, with 90% or more of the sample endorsing the most positive scale value for these three items. With respect to computer anxiety, 9% reported great or major anxiety (8 of 93), while 20% indicated some anxiety (19 of 93).

33

Two items served as overall indicators of student reactions. One asked for a comparison with the paper-pencil format and the other for a global evaluation of the day's experience. Concerning the overall experience of the day's testing, the mean on a 5-point scale was 3.87 (SD=1.22). Specifically, 69 of 92 (75%) students reported that Internet testing was good or great, 16 (17%) reported a neutral evaluation, 7 (8%) reported that Internet testing was satisfactory but needed work, and only 1 (1%) student indicated that Internet testing was a disaster. Concerning a comparison of paper-pencil and Internet formats, the mean score was 3.92 (SD=.89). Specifically, 58 of 93 (62%) students indicated that Internet testing was better or much better than the traditional format, 22 (23%) reported a neutral opinion, 8 (9%) reported that Internet testing was satisfactory but preferred paper-pencil, and 5 (6%) indicated that paper-pencil tests were much better. However, the reader should note that only 44 of the 61 individuals with scores on both assessments responded to this question. Finally, the relationship between test scores, on the one hand, and overall reaction and format comparison, on the other, were .42 and .34. These results, taken together, indicate a positive overall evaluation of the testing experience and a positive comparison of the Internet format relative to the traditional format. Attitudinal reactions were positive for the Internet format (although we did not obtain focus group or interview data with students to explore their specific concerns).

The fourth research objective, to evaluate technical feasibility, was addressed through telephone interviews with teachers. The logistical and technical issues surrounding implementation were of specific interest. Thus far, two occupational areas, Automotive Body Repair and Child Care/Guidance, have been investigated in addition to Administrative Office Technology. A comparison between formats revealed slight performance differences. The reactions of teachers examined by telephone interview were mostly positive. The major problems teachers reported during the interviews were related to the technical details of Internet assessment and to some problems with presentation and scrolling of items on the screen (e.g., items with graphics).

CONCLUSIONS AND FUTURE DIRECTIONS

Alternative assessment formats and techniques are assuming greater importance as educational institutions position themselves to take advantage of technology. As evidence of this, a special feature of the journal *Techniques* (March, 1998), entitled "Putting Assessment to the Test" reviewed five innovations in assessment. The innovations were titled "Is there a best way to test?, Engineering a grade, Windows on progress, Show & tell, and Online testing. The current project, which evaluated occupation-specific assessment over the Internet was successful as measured by several quantitative and qualitative indicators. The direct comparison (delayed because of a time lag in receiving the scored test data from a contractor, which itself supports the Internet testing concept) indicated that the Internet test scores were higher than the paper-pencil version, which was expected because the same form of the test was used. Scores on the Internet assessments were comparable to previous years in terms of average, variability, and internal consistency. The attitude survey responses indicated very positive evaluations of the process of Internet testing and correspondingly positive reactions to the overall experience and preferences for the Internet format. Qualitatively, teachers/facilitators reported that they and the students were impressed with the rapidity of reporting of results via an overall score and scores on six scoring clusters (ranging from office equipment/procedures to professionalism). Their positive reaction is easy to understand, given a two-week delay in receiving score reports from the contractor. Although the delay is exacerbated by school districts that are late in returning completed assessments, incorrect entries on answer documents, and the contractor's need to produce reports, it can be frustrating for teachers and students.

Based on the results of this small pilot study, it appears that there is potential for future testing via the Internet. The policy and planning steps required for implementation at the state level should be carefully considered, but would depend on the overall strategy determined by the state's Department of Education. Zakrzewski and Bull (1998) used their experiences over a four-year period to provide guidance for institutional strategies when considering implementation. Their framework involves a consideration of three areas of possible contribution from technology: formative assessment, summative assessment, and computer-aided learning. They advocate using the progression from summative assessment to formative assessment to computer-aided learning.

Whatever the framework, many of the decisions are the province of the State's Department of Education. One strategy would involve a gradual rollout across the state. One example would be to extend the Internet testing concept to the remaining Administrative Office Technology programs. One advantage of this tactic would be that the AOT curriculum makes extensive use of computers. Another tactic would be to extend the Internet testing concept to an entire cluster, say Agricultural Education, and then to proceed cluster by cluster to roll out the program. Whatever the strategy, one important preliminary component should be a technology assessment to ensure that targeted Vocational Education Planning Districts (VEPD) possess the hardware and software required for computerized systems. Specifically, the software used in the present project was developed using a philosophy of the lowest common denominator, thus Internet access using browsers (e.g., Netscape) is the basic requirement. In considering the task of moving the state's vocational education programs toward Internet testing, three interrelated dimensions seem of primary interest for implementation and marketing. In order of importance as they appear to us, the dimensions are costs and benefits to stakeholders, details of the implementation, and evaluation of effectiveness of the computerized testing modality.

Further, it should be clear that we are not recommending that all assessment in vocational education move immediately to a computer adaptive testing (CAT) format, or even to a computerized format. While there are advantages to CAT (Wilson, Genco, & Yager, 1985), there are still significant advantages in using the computer as a page-turner that do not require advanced psychometric theory, computer programming, or huge empirically-validated item pools. Specifically, the speed of feedback benefits both learners and instructors, the capability to assess later in the school year benefits those same two groups, and the reduction in printing and shipping costs benefits administrators and test developers. For example, software that offers most of the features of the software system that we used, specifically database capability for test items and Internet server interface with security features, is available at fairly modest prices. Low cost makes the technology more readily available to many vocational education systems. A second option for those schools that lack Internet access is to provide a self-contained assessment system on a compact disk (CD) that can be taken to the schools. Given the low cost of creating CDs (e.g., using CD-R and CD-RW peripherals), cost is not an issue. Also, it should be remembered that there are some dimensions of the career-technical curriculum that computers cannot capture well at present, for example psychomotor skills. However, psychomotor dimensions are not impossible to assess but merely more difficult to program and might also require additional equipment for responding. For example, we maintain that the use of digital drawing tablets would permit the assessment of manual dexterity (i.e., in personnel selection) or artistic renderings (i.e., for graphic artists).

Results of this pilot project, similar to those of the Pennsylvania project, indicate that Internet testing performs comparably and is preferred by students over the traditional paper-pencil format.

Even though the present usage of the computer (page-turner) corresponds to what Hambleton (1996) termed 'the earlier use of computers in assessment,' it appears that this usage can be efficient and effective. Thus, we suggest that the advantages of Internet testing outweigh the disadvantages and that continued experimentation with this medium will create improvements that resolve those problems. In sum, this medium can migrate from "vision to reality" (Bicanich et al., 1997).

REFERENCES

Bicanich, E., Hardwicke, S.B., Slivinski, T., & Kapes, J.T. (1997). Internet-based testing: A vision or reality? *Technological Horizons in Education, 24,* 61-65.

Drasgow, F., & Olson-Buchanan, J.B. (Eds.). (1999). *Innovations in computerized assessment.* Mahwah, NJ: Erlbaum.

Drasgow, F., Olson, J.B., Keenan, P.A., Moberg, P., & Mead A.D. (1993). Computerized assessment. In G.R. Ferris & K.M. Rowland (Eds.), *Research in personnel and human resources management* (pp. 156-202). Greenwich, CT: JAI Press.

Feldman, S.E., & Krumenaker, L. (1995). *The Internet at a glance: Finding information on the Internet* (3rd ed.). Medford, NJ : Information Today.

Hambleton, R.K. (1989). Principles and selected applications of item response theory. In R.L. Linn (Ed.), *Educational measurement* (pp. 147-200). New York: American Council on Education/Macmillan.

Hambleton, R.K. (1996). Advances in assessment models, methods, and practices. In D.C. Berliner & R.C. Calfee (Eds.), *Handbook of educational psychology* (pp. 889-925). New York: American Council on Education/Macmillan.

Hahn, H., & Stout, R. (1994). *The Internet complete reference.* Berkeley, CA: Osborne McGraw-Hill.

Institute for Information Studies (1997). *The Internet as paradigm.* Falls Church, VA: Author.

Kapes, J.T., Martinez, L., Ip, C-F., Slivinski, T., & Hardwicke, S. (1998). Internet-based vs. paper-pencil occupational competency test administration: An equivalency study. *Journal of Vocational Education Research, 23,* 201-219.

Kingsbury, G.G., & Houser, R.L. (1999). Developing computerized adaptive tests for school children. In F. Drasgow & J.B. Olson-Buchanan (Eds.), *Innovations in computerized assessment* (pp. 93-115). Mahwah, NJ: Erlbaum.

Lee, J.A., Moreno, K.E., & Sympson, J.B. (1986). The effects of mode of test administration on test performance. *Educational and Psychological Measurement, 46,* 467-474.

Maran, R. (1995). *Internet and World Wide Web simplified.* Foster City, CA: IDG Books.

New Forms of Assessment [Special issue]. *Techniques, 73*(3).

Resnick, L.B., & Wirt, J.G. (Eds.). (1996). *Linking school and work: Roles for standards and assessments.* San Francisco, CA: Jossey-Bass

Sands, W.A., Waters, B.K., & McBride, J.R. (Eds). (1997). *Computerized adaptive testing: From inquiry to operation.* Washington, DC: American Psychological Association.

Slivinski, T., Hardwicke, S.B., Kapes, J.T., Boyer, E., Ip, C-F., & Martinez, L. (1997, July). *Pennsylvania–VTECS Internet testing pilot project* (Technical Report submitted to Pennsylvania Department of Education). Arlington, VA: WebTester, Inc.

Stephens, D., Bull, J., & Wade, W. (1998). Computer-assisted assessment: Suggested guidelines for an institutional strategy. *Assessment & Evaluation in Higher Education, 23,* 283-294.

Thomas, B.J. (1995). *The Internet for scientists and engineers: Online tools and resources.* Bellingham, WA: SPIE Optical Engineering Press.

Wilson, F.R., Genco, K.T., & Yager, G.G. (1985). Assessing the equivalence of paper-and-pencil vs computerized tests: Demonstration of a promising methodology. *Computers in Human Behavior, 1,* 265-275.

Zakrzewski, S., & Bull, J. (1998). The mass implementation and evaluation of computer-based assessments. *Assessment & Evaluation in Higher Education, 23,* 141-52.

Ohio Pilot Project Evaluation: Thank you very much for taking this Internet-based test of your occupational specialty. In order to help us evaluate and improve our computer-based testing, please answer each of the questions below about your test experience today. After you have completed this evaluation, please push the SEND EVALUATION button at the bottom of this form.

1. If you were assisted in your test by a facilitator, check this bullet.   ○
   If no facilitator, check this bullet.   ○

PRIOR COMPUTER AND INTERNET EXPERIENCE

2. Do you have a computer at home?
   Yes, I have a computer at home.
   No, I do not have a computer at home.
3. Do you have daily access to a computer at school?
   Yes, I have daily access to a computer at school.
   No, I do not have daily use of a computer at school.
4. How would you rate your prior computer experience?
   No prior experience.
   Little prior experience.
   Moderate experience.
   Lots of experience.
   A real computer pro.
5. How would you rate your prior Internet experience?
   No prior experience.
   Little prior experience.
   Moderate experience.
   Lots of experience.
   Real computer pro.

TODAY'S INTERNET TEST EXPERIENCE.

6. Did you have any problems accessing our Website or getting your test downloaded
   Major problems getting the test
   Some minor problems accessing or downloading the test
   No problems
7. Did you have any problems using the mouse to answer the questions?
   Major problems using the mouse
   Some minor problems using the mouse
   No problems
8. Did you have any problems using the mouse to move from one question to the next?
   Major problems using the mouse
   Some minor problems using the mouse
   No problems
9. Did you have any problems reading the questions on the computer screen?
   Major problems reading the questions
   Some minor problems reading the questions
   No problems

10. Did you have any problems submitting your answers and getting your test scored?
    Major problems with the scoring
    Some minor problems with the scoring
    No problems
11. Did you have any problems reading your test scores or printing them out?
    Major problems saving or printing test results
    Some minor problems saving or printing the results
    No problems
12. To what extent did you have problems with keyboarding?
    Major problems with keyboarding
    Great problems with keyboarding
    Some problems with keyboarding
    Minor problems with keyboarding
    No problems
13. Please look at your TEST RESULTS. Do you understand what each section means and how well you scored on the test?
    I do not understand the results
    I understand my basic percentage correct score, but do not understand the other results
    I understand my basic percentage correct score and overall results, but not my results by subject
    understand most of the results
    I understand all of the results
14. To what extent were you anxious while working with computer technology?
    Great anxieties
    Major anxieties
    Some anxiety
    Little anxiety
    No anxiety

COMPARE TODAY'S INTERNET TEST WITH OTHER PAPER-PENCIL TESTS

15. How do you rate your overall experience today using the Internet compared to the usual paper-pencil test?
    Paper-pencil tests are much better
    Although I like paper-pencil tests, the Internet test was OK
    It makes no difference to me
    The Internet test is better than the paper-pencil tests
    The Internet test was much better than paper-pencil tests
16. Overall, how do you rate your experience today using the computer and Internet for testing?
    It was a disaster
    It was OK, although Internet testing needs a lot of work
    So-so
    The Internet testing was good
    The Internet testing was great