

TESTING IN A COMPUTER TECHNOLOGY COURSE: AN INVESTIGATION OF EQUIVALENCY IN PERFORMANCE BETWEEN ONLINE AND PAPER AND PENCIL METHODS

Melody W. Alexander

Ball State University

James E. Bartlett

University of Illinois

Allen D. Truell

Karen Ouwenga

Ball State University

ABSTRACT

This experiment sought to examine the equivalence of online and paper and pencil testing methods as related to student performance in a computer technology course. Test score and completion time were the dependent variables that were used to assess students' performance. The study utilized a quasi-experimental design. Test scores were not significantly different on the variables of pretest, age, class standing, ethnicity, and gender. The findings showed that test scores were equivalent in both groups; however, time to complete the test was significantly different between the groups. The online testing group completed the test in less time than the paper and pencil group. The exploration of class standing did reveal that freshmen were the only group that took significantly less time to complete the online test. The study supports the online test method did not effect score as result of age, class level, and gender.

With high demands on curriculum coverage within the classroom, career and technical education teachers are in need of an efficient method to conduct assessment activities without lessening their impact or purpose. Test administration is one type of activity that can be proctored. The integration of technology into the classroom is now affordable and realistic for most educational institutions. One of the latest technological advances that has potential to impact education is online testing.

In the 1980s, the introduction of the personal computer caused an excitement in education that has yet to be paralleled (Miller, 2000). Within the realm of education, computers assumed supportive roles in teaching and learning (Gibson, Brewer, Dholakia, Vouk, & Bitzer, 2000; Miller; Newby & Fisher, 1998). Career and technical education teachers can use video clips, sound bites, animated graphics, photographs, tables and graphs, drawings, special effects, and more recently, the Internet to enhance instruction (Basics of Computer-Based Testing and Assessment, 2000; Doughty, Magill, & Turner, 1996; Hazari, 1998; MacDougall, Place, & Currie, 1998; Song, 1998).

Multimedia and hypermedia, use of multiple forms of media mixed with technology in conjunction with a microcomputer, distance learning, distance education, and traditional classroom supportive materials has taken on a whole new image (Havice, 2000; Thomson & Stringer, 1998). Miller (2000) found that the introduction of computers into instruction increased the amount of learning in a shorter amount of time and overall has improved students' attitudes towards education. Furthermore, the impact of technology in the delivery of instruction has reduced barriers of time and distance for students (Song, 1998).

Along with distance education comes the experience of student assessment in a non-traditional format. Students now submit course work by e-mail, complete learning activities through the World Wide Web, and complete student assessments in the form of online testing (Basics of Computer-Based Testing and Assessment, 2000; Bishop, 2000; Chauncey, 1995; Doughty et al., 1996; Gibson et al., 2000; Hazari, 1998; Newby & Fisher, 1998; Newman, 2000; Shermis & Lombard, 1998; Thomson & Stringer, 1998; Treadway, 1997). Online testing is typically seen in the form of a database of multiple choice questions posted on the Internet with secured access (Bocij & Greasley, 1999; Bull, 1996; Daly, 2000; Doughty et al.; Hazari; Greenberg, 1998; Gibson et al.; Kumar, 1996; Treadway, 1997,1998; Zakrzewski & Bull, 1998). Even though multiple choice questions are the typical form of assessment seen on the Internet, many software programs also have the capability of using fill-in-the-blank, matching, and essay questions, and some are even capable of producing tests that use a variety of multimedia tools (Basics of Computer-Based Testing and Assessment; Chauncey; Doughty et al.; Hazari; Judge, 1999; Thomson & Stringer).

There are concerns with the use of online testing methods for student assessment. One concern is the lack of resources; more specifically, the limited hardware, software, and technical expertise that may be needed (Basics of Computer-Based Testing and Assessment, 2000; Bishop, 2000; Bull, 1996; Newby & Fisher, 1998; Zakrzewski & Bull, 1998). A second concern lies in the area of security and reliability of the testing system (Bishop; Bull; Zakrzewski & Bull). An additional system, or a back-up plan, should be in place in the event of a breakdown of the system. Teachers also need to be insured that students who are getting credit of the assessments are the ones completing the online test. Finally, there is an overall concern that online testing will have either positive or negative effects on student test scores when compared with traditional testing methods (Bocij & Greasley, 1999). Furthermore, educational researchers are concerned if other variables (gender, special education needs, economic/educational backgrounds, or disabilities) place sub groups at disadvantages when measuring achievement (Bicanich, Slivinski, Hardwicke, & Kapes, 1997).

Even though there are some concerns in the area of online testing, there are many positive features. One benefit is that tests can be scheduled when it is convenient for the student, which also encourages students to increase time management skills (Basics of Computer-Based Testing and Assessment, 2000; Cochran, 1998; Greenberg, 1998; Judge, 1999; Song, 1998). Computer-based tests taken online can be scored immediately, which means students are able to receive feedback within a matter of seconds (Basics of Computer-Based Testing and Assessment; Bishop, 2000; Cochran; Daly, 2000; Gibson et al., 2000; Gokhale, 1996; Greenberg; Judge; Song; Thomson & Stringer, 1998). After the tests are scored, the data can be easily downloaded into an electronic gradebook system for teacher convenience (Cochran; Greenberg; Treadway, 1997, 1998).

Another major benefit of online testing is the amount of time that is saved compared to the traditional paper and pencil test (Bocij & Greasley, 1999; Greenberg, 1998; Newman, 2000; Shermis & Lombard, 1998; Song, 1998). Since the paper tests are no longer needed, institutions are able to save money that would have been spent on the paper for the exams, and the time spent to score the exams (Newman; Song).

There are many benefits for using online testing. Approximately 10% of high schools and 30% of universities in the United States have established computer labs specifically for online testing (Greenberg, 1998). However, there are some gray areas in computer-based testing that still should be explored before its true effectiveness is known. In a pilot of an online testing program with high school vocational students in Pennsylvania, results appeared to be equivalent with traditional tests and bias related to gender, educational needs, and economic status were not present (Bicanich et al., 1997). Although the literature is clear that online testing saves time, it is not clear if online testing results are equivalent with traditional testing results. Student demographic characteristic such as gender, age, and year in school, were studied because they have been shown to be explanatory in student performance (Agarwal & Day, 1998). The present study, therefore, will compare the variables of student achievement as measured by grade and student performance as measured by time to complete the assessment with the online testing and traditional paper and pencil groups.

NEED FOR THE STUDY

Although online testing is a technology most educational institutions will be able to implement, research is lacking in identifying the affect this type of testing has on performance specifically measured by grade and time to complete the assessment. A comparison of traditional test taking results with online test results would be helpful for career and technical educators as they begin to consider implementing this new technological activity into their courses.

STATEMENT OF THE PROBLEM

Technology has led to many changes in the classroom. It is necessary, however, to ensure that these changes are positive. Thus, the problem was to examine if differences in student performance exists in terms of test score and time to complete assessments using traditional and online methods. To investigate this problem, an quasi-experiment was conducted using exam grades from students at a mid-sized, Midwestern state university. The following research questions were addressed:

1. Is there a statistically significant difference between online testing and traditional paper and pencil test scores?
2. Is there a statistically significant difference between online testing and traditional paper and pencil time to complete test?
3. Are there statistically significant relationships between the time it takes to complete an online and traditional paper and pencil test and the score?
4. Are there statistically significant differences between online testing and traditional paper and pencil test scores in relation to the selected demographic variables of age, class standing, and gender?

5. Are there statistically significant differences between online testing and traditional paper and pencil time to complete tests in relation to the selected demographic variables of age, class standing, and gender?

PURPOSE OF THE STUDY

This study compared differences between online testing and traditional paper and pencil testing methods in relation to grades, test time, and demographic differences. The study results will provide educators, administrators, and curriculum planners with documentation to make decisions in regard to using or not using online testing in their courses.

METHODOLOGY

A quasi-experimental design was used to control for as many threats to internal validity as possible. This design was used due to the use of intact groups and the lack of ability to have randomization. A pretest-posttest design was used (Campbell & Stanley, 1966). To control for the testing effect, the main concern with this design, the pretest instrument only had a random sample of questions from the posttest.

PARTICIPANTS

Two intact classes of college students from a course in the business education department at a mid-west research intensive university were selected to participate in this project. The study population consisted of two sections of a business technology course with a total of 79 students (40 in traditional group and 43 in online testing group). The business technology course covered introductory to computer theory concepts and computer applications programs including word processing, spreadsheet, and database. This group of students was a purposeful sample to examine students in technology courses (Gall, Borg, & Gall, 1996).

PROCEDURES OF THE STUDY

Each class used the same course materials (book, software, handouts, etc.), received the same lecture by the same instructor, and completed the same projects. A written pretest was given to all students to determine the content knowledge achievement for the specific unit. The posttest was administered to one group in a traditional paper and pencil method using scantrons (control group), and the other group took the posttest using an online testing method in a proctored lab (experimental group). Both the pretest and posttest were examined for validity by three experts in the course content area.

Each group was given the same pretest to establish equivalent groups. After the pretest, the same lessons were given to both groups and the same topics and objectives covered. One class was administered a theory test in the traditional paper and pencil method, while the remaining class took the test online in a proctored computer lab. The exact same questions were used, and the time allotment was 30 minutes for both sections. Following the procedures approved by the Institutional Review Board, after the test scores were recorded for grading purposes, any and all identifiers were removed before statistical analysis began.

DATA ANALYSIS

Data were analyzed using frequencies and percentages as appropriate to describe participants. To identify if any significant differences existed between test scores and time of test completion

between online testing and traditional testing groups, *t*-tests were used. Pearson's product moment coefficient was used to determine the relationship between test score and time of test completion. ANOVA was used to determine any significant differences between test scores and test time in relation to the demographic variables of gender and rank in class, and ANCOVA was used to determine any significant differences between test scores and test time in relation to age. Orthogonal contrasts were used to determine if significant difference existed between time and rank in class. Significance was set a priori at the .05 level.

FINDINGS

The analysis of the findings of this study identified: (a) differences between online testing and traditional paper and pencil test scores, (b) differences between online testing and traditional paper and pencil time to complete test, (c) relationships between the time it takes to complete an online and traditional paper and pencil test and the score, (d) differences between online testing and traditional paper and pencil test scores in relation to the demographic variables of gender, age, or class standing; and (e) differences between online testing and traditional paper and pencil time to complete tests in relation to the demographic variables of gender, age, or class standing.

DEMOGRAPHIC PROFILE OF PARTICIPANTS

The first step in the investigation was to provide evidence the groups were equal. In order to accomplish this, 10 questions were administered as a pretest to each group. On the posttest, the paper and pencil group scored an average of 53.2%, while the online group scored an average of 49.8%. This provided evidence there was no significant difference ($p = .94$) between the groups. The average participant was 20.27 ($sd = 1.41$) years old, with age range of 18-25. The demographic breakdown of the two study groups can be seen in Table 1. An analysis of the demographic variables between the online and traditional groups revealed no significant differences.

Table 1

Demographic Profile of Participants

Factor	Participants (N = 79)	
	f	%
Gender		
Male	43	54.0
Female	36	46.0
Ethnicity ^a		
Caucasian	66	89.2
African American	5	6.8
Hispanic	3	4.1
Class Standing		
Freshman	24	30.4
Sophomore	34	43.0
Junior	12	15.2
Senior/Graduate	9	11.4

NOTE. ^aSome participants chose not to disclose ethnicity information.

COMPARISON BETWEEN ONLINE AND PAPER AND PENCIL TESTS BY TEST SCORE AND TIME

Research questions one and two sought to explore if there were any significant differences between online and paper and pencil testing methods in relation to test score and test time. The analysis of the test scores and time taken for the exam is displayed in Table 2. The mean scores of traditional test was 22.03 ($sd = 2.77$) which is a 73%, and the mean score of the online test was 22.60 ($sd = 2.77$) which is a 77%. The test scores showed no significant difference between the two groups. However, there was a significant difference ($p = .02$) in the time used to take the exam. Participants who took the exam using the online testing method completed the test significantly faster than those using the paper and pencil method.

Table 2

Comparison Online and Paper and Pencil Testing Methods with Test Grade and Time

Variable	<u>M</u>	<u>sd</u>	<u>t</u>	<u>df</u>	<u>p</u>
Test Score					
Online	22.60	2.77	.884	77	.380
Paper and Pencil	22.03	3.03			
Test Time					
Online	10.80	3.49	-2.353	77	.021*
Paper and Pencil	12.52	2.94			

Note. *Significance at the .05 level.

COMPARISON BETWEEN ONLINE AND PAPER AND PENCIL TEST SCORE AND TIME WITH DEMOGRAPHIC VARIABLES

Research question three sought to explore if relationships existed between the time it took to complete an online and paper and pencil test and the score. Table 3 shows that a moderate correlation ($r = .359$, $p = .03$) existed in the traditional group and a negligible correlation existed in the online group.

Table 3

Relationship Between Test Score and Time for Online and Paper and Pencil Testing Groups

Score	<u>r</u>	Time Interpretation	<u>p</u>
Paper and Pencil Group	.359	Moderate	.03
Online Group	.081	Negligible	.61

Note. Interpretations according to Davis' (1971) descriptors: .01-.09 (negligible), .10-.29 (low), .30-.49 (moderate), .50-.69 (substantial), .70-.99 (very high), and 1.0 (perfect)

Research question four examined if significant differences existed between online and paper and pencil test scores in relationship to demographic variables of age, rank in class, ethnicity, and

gender. Ethnicity was not compared due to the low size in the experiment. Analysis of variance of score by gender and testing treatment showed no significant differences. Analysis of covariance revealed score was not significantly different in relationship to age and testing method. In addition, analysis of variance of score by rank in class and treatment found no significant differences existed. Table 4 illustrates demographic comparisons related to score.

Table 4

Analysis of Variance and Analysis of Covariance of Score by Demographics

Analysis of Variance of Score by Gender and Treatment

	SS	df	MS	F	p
Intercept	38260.7	1	38260.73	4552.27	<.01
Treatment	8.88	1	8.88	1.06	.31
Gender	4.6	1	4.60	0.55	.46
Treatment * Gender	9.53	1	9.53	1.13	.29
Error	630.35	75			
Total	40083	79			

Analysis of Covariance of Score by Age and Treatment

	SS	df	MS	F	p
Intercept	198.61	1	198.61	23.47	<.01
Age	0.02	1	.02	.01	.96
Treatment	6.52	1	6.52	.77	.38
Error	643.23	76			
Total	40083	79			

Analysis of Variance of Score by Treatment and Class

	SS	df	MS	F	p
Between Groups	60.02	7	8.58	1.03	.42
Within Groups	589.75	71	8.31		
Total	649.77	78			

Research question five examined if significant differences existed between online and paper and pencil time for test completion in relationship to demographic variables of age, class standing, ethnicity, and gender. Ethnicity was not compared due to the low size in the experiment. Analysis of variance of time by gender and testing treatment showed only significant difference in treatment method. Analysis of covariance revealed time was not significantly different in relationship to age; however, as in the previous analysis, the testing method was significant. In addition, analysis of variance of time by rank in class and treatment found significant differences existed. Comparisons were pre-planned if significant differences existed. Table 4 illustrates demographic comparisons related to score. The orthogonal contrasted revealed a significant difference appeared between the freshman class and time it took to complete the online or paper pencil test. All other ranks in class were not significantly different as shown in Table 6.

Table 5

Analysis of Variance and Analysis of Covariance of Time by DemographicsAnalysis of Variance of Time by Gender and Treatment

	SS	df	MS	F	p
Intercept	10493.8	1	10493.8	977.47	<.01
Treatment	56.33	1	56.33	5.34	.03
Gender	2.43	1	2.43	.23	.64
Treatment * Gender	3.98	1	3.98	.37	.55
Error	805.17	75			
Total	11473.12	79			

Analysis of Covariance of Time by Age and Treatment

	SS	df	MS	F	p
Intercept	109.69	1	109.69	10.39	<.01
Age	10.21	1	10.21	.97	.33
Treatment	58.31	1	58.31	5.53	.02
Error	802.08	76			
Total	11473.12	79			

Analysis of Variance of Time by Treatment/Class

	SS	df	MS	F	p
Between Groups	213.97	7	30.57	3.31	<.01
Within Groups	656.73	71	9.25		
Total	870.7	78			

Table 6

Orthogonal Contrasts to Show the Comparisons of Time by Class Level

	Online Mean	SD	Traditional Mean	SD	df	t	p
Freshmen	13.8	2.99	8.9	2.10	71	3.91	<.001
Sophomores	12.1	2.64	12.7	4.06	71	-.55	.59
Juniors	11.0	2.53	10.5	1.85	71	.30	.77
Seniors	13.0	5.66	9.5	2.72	71	1.45	.15

CONCLUSIONS

Research questions one and two identified any significant differences between online and paper and pencil testing methods in relation to test score and time. Results from this study indicated that taking an exam online as compared to the tradition paper and pencil testing does not have an effect on overall exam scores. However, there is a savings in time between testing methods for students. Online tests take significantly less time to complete than paper and pencil tests.

Research question three examined the relationship between test score and the time it took to complete the test. Online scores did not significantly relate with the time to complete the test. However, paper and pencil scores did significantly relate with the time to complete the test.

Research questions four and five compared online and paper and pencil test scores and time with the demographic variables of age, gender, and rank in class. As no significant differences were found in score, it is likely that demographic variables do not have an effect on online or paper and pencil testing methods in relation to achievement level on exams. Time, however, did reveal a significant difference for the treatment and specifically for rank in class. Freshmen took less time on the online test than on the traditional paper and pencil test. This difference needs to be examined further.

DISCUSSION

From the score and time analysis, it is evident that online testing is more efficient for students in relationship to time. This finding supports previous findings of time-saving measures (Bocij & Greasley, 1999; Greenberg, 1998; Newman, 2000; Shermis & Lombard, 1998; Song, 1998). A major concern when switching testing methods focuses on student achievement. The data gathered and analyzed showed that online testing could be used without sacrificing student scores. These findings also support the experiment Bicanich et al. (1997) conducted with high school vocational students that score is not different among gender and provides evidence these results are similar with college-level students.

Online testing time was not shown to correlate with test score, as did the traditional testing method. This finding may alleviate the concerns that students who have more time to complete an exam do better. Online testing could play a major part in all levels of postsecondary education. Specifically, freshmen took less time to complete the online test and achieved similar scores. This also supports Agarwal and Day (1998) who suggested individual characteristics explain variance in student performance. However, the common concern of varying test scores and unfair advantages when changing testing methods (Bocij & Greasley, 1999) was not supported in this study. With testing times greatly reduced, a teacher would not necessarily need to sacrifice an entire class period for testing alone, and students would achieve equivalent results. With the heavy emphasis on standards, any extra time could play an important role in the student's learning experience.

RECOMMENDATIONS FOR FURTHER STUDY

The following are recommendations for further research and study in the area of online testing and its role in education:

1. As this study focused on comparing online and paper and pencil testing methods, further research should be conducted to measure students' attitudes and perceptions towards online testing. This would provide a beginning to examine how students view online testing methods.

2. As this study focused on student outcomes, future studies should be conducted to identify the use of online testing by teachers, as well as to measure the time saved by teachers in the overall grading and evaluation of test scores, comparing online with paper and pencil testing methods. This type of study could provide evidence to the amount of performance that could be improved through the implementation of online testing.
3. As technology changes so rapidly, further research should compare new testing methods as they emerge. This research would provide support that assessment of students is not biased by unchangeable demographic variables.

REFERENCES

- Agarwal, R., & Day, E. A. (1998). The impact of the internet on economic education. *Journal of Economic Education*, 29(2), 99-115.
- Basics of computer-based testing and assessment, The* (2000). Retrieved from <http://www.jmu.edu/assessment/cbtbasics.htm>
- Bicanich, E., Slivinski, T., Hardwicke, S., & Kapes, J. (1997). Internet-based testing: A vision or reality? *THE Journal*, 25(2), 61-65.
- Bishop, P. (2000). CTI maths and stats workshop 18. Retrieved from <http://www.bham.ac.uk/ctimath/workshops/wass.htm>
- Bocij, P., & Greasley, A. (1999). Can computer-based testing achieve quality and efficiency in assessment? *International Journal of Educational Technology*, 1(1), 1-18. Retrieved from <http://www.outreach.uiuc.edu/ijet/v1n1/bocij/index.html>
- Bull, J. (1996). Computer based assessment: Some issues for consideration. Retrieved from <http://www.cti.ac.uk/publ/actlea/issue1/bull/>
- Campbell, G. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Chauncey, H., Jr. (1995). A calm before the storm? *Yale Alumni Magazine*, 58(7), 30-31.
- Cochran, E. P. (1998, March-April). The mouse replaces the pencil: TOEFL goes electronic. *ESL Magazine*, 1(2), 10-12.
- Davis, J. A. (1971). *Elementary survey analysis*. Englewood Cliffs, NJ: Prentice-Hall.
- Daly, T. (2000). Computer based assessment (CBA). Retrieved from <http://www.mcc.ac.uk/newsletters/Local/issue72/cba.html>
- Doughty, G., Magill, J., & Turner, I. (1996). Interactive multimedia computer based assessment. Retrieved from <http://www.cti.ac.uk/publ/actlea/issue1/doughty/>
- Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational research: An introduction* (6th ed.). White Plains, NY: Longman.

Gibson, E. J., Brewer, P. W., Dholakia, A., Vouk, M. A., & Bitzer, D.L. (2000). A comparative analysis of web-based testing and evaluation systems. Retrieved from <http://renoir.csc.ncsu.edu/MRA/Reports/WebBasedTesting.html>

Gokhale, A. A. (1996). Effectiveness of computer simulation for enhancing higher order thinking. *Journal of Industrial Teacher Education*, 33(4), 36-46.

Greenberg, R. (1998). Online testing. *Techniques*, 73(3), 26-28. Retrieved from <http://47.226.94.254:5239/per?sp.nextform=fullrec.htm&sp.usernumber.p=459434>

Havice, W. L. (2000). College students' attitudes toward oral lectures and integrated media presentations. Retrieved from <http://scholar.lib.vt.edu/ejournals/JTS/Winter-Spring-1999/PDF/havice.pdf>

Hazari, S. (1998). *Online testing methods for web courses*. Presented at the 1998 Distance Teaching and Learning Conference. (ERIC Document Reproduction No. ED422835)

Judge, G. (1999). The production and use of online web quizzes for economics. *Computers in Higher Education Economics Review*, 13(1). Retrieved from http://www.ilrt.bris.ac.uk/ctiecon/cheer/ch13_1/ch13_1p21.htm

Kumar, D. (1996). *Computers and assessment in science education*. (ERIC Document Reproduction No. ED395770)

MacDougall, G., Place, C., & Currie, D. (1998, June). *Web-based testing: A form-based template for creating multimedia enhanced tests*. Paper presented at the 1998 World Conference on Educational Multimedia and Hypermedia and World Conference on Educational Telecommunications, Freiburg, Germany. (ERIC Document Reproduction No. ED428692)

Miller, L. W. (2000). Computer integration by vocational teacher educators. *Journal of Vocational and Technical Education*, 14(1). Retrieved from <http://scholar.lib.vt.edu/ejournals/JVTE/v14n1/JBTE-3.html>

Newby, M., & Fisher, D. (1998). *The association between computer laboratory environment and student outcomes*. Paper presented at the Australian Association for Research in Education Annual Conference, Adelaide, Australia. Retrieved from <http://www.swin.edu.au/aare/98pap/new98037.html>

Newman, C. (2000). Online testing rated. *Advertising-Age*, 71(20), 64. Retrieved from <http://147.226.94.254:5239/per>

Shermis, M. D., & Lombard, D. (1998). *Effects of computer-based test administrations on test anxiety and performance*. (ERIC Document Reproduction No. EJ561400)

Song, J. K. (1998). *Using the world wide web in education and training*. Presented at 1998 Information Technology in Education and Training Conference – session 1. (ERIC Document Reproduction No. ED417703)

Thomson, J. S., & Stringer, S. B. (1998, August). *Evaluating for distance learning: Feedback from students and faculty*. Paper presented at the Annual Conference on Distance Teaching and Learning, Madison, WI. (ERIC Document Reproduction No. ED422835)

Treadway, R. (1997, June). *Integrating a computerized testing system and electronic lecture notes in first-year mathematics courses*. Paper presented at the Association of Small Computer Users in Education Summer Conference Proceedings, North Myrtle Beach, SC. (ERIC Document Reproduction No. ED410938)

Treadway, R. (1998, June). *An integrated computerized instructional system for classroom and lab*. Paper presented at the Association of Small Computer Users in Education: Proceedings of the ASCUE Summer Conference, North Myrtle Beach, SC. (ERIC Document Reproduction No. ED425736)

Zakrzewski, S., & Bull, J. (1998). The mass implementation and evaluation of computer based assessments. *Assessment and Evaluation*, 23(2), 141-152.