DOCUMENT RESUME

ED 476 992 IR 021 720

AUTHOR	Davies, JoAnne E.; Szabo, Michael; Montgomerie, Craig
TITLE	Assessing Information and Communication Technology Literacy of Education Undergraduates: Instrument Development.
PUB DATE	2002-06-00
NOTE	8p.; In: ED-MEDIA 2002 World Conference on Educational Multimedia, Hypermedia & Telecommunications. Proceedings (14th, Denver, Colorado, June 24-29, 2002); see IR 021 687.
AVAILABLE FROM	Association for the Advancement of Computing in Education (AACE), P.O. Box 3728, Norfolk, VA 23514. Tel: 757-623-7588; e-mail: info@aace.org; Web site: http://www.aace.org/DL/.
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	*Computer Literacy; Computer Mediated Communication; *Computer System Design; Computer Uses in Education; *Educational Assessment; *Educational Technology; Higher Education; Information Technology; *Teacher Education; Teacher Education Programs; *Technology Integration

ABSTRACT

In recent years, the view that Information and Communication Technology (ICT) is vital in K-12 education has become widespread. ICT use in schools has increased and various professional bodies have set ICT standards for students and teachers. Schools of education are under pressure to produce teachers who are able to effectively integrate technology into their teaching. However, most teacher preparation programs do not adequately prepare teachers in ICT, nor assess candidates relative to ICT standards. This paper discusses the development of a computerized system to assess ICT declarative and procedural knowledge and to provide a profile to the participant. (Contains 12 references and 1 table.) (Author)

Assessing Information and Communication Technology Literacy of Education Undergraduates: Instrument Development

JoAnne E. Davies <joanne.davies@ualberta.ca>
Michael Szabo <mike.szabo@ualberta.ca>
Craig Montgomerie <craig.Montgomerie@ualberta.ca>
Department of Educational Psychology
University of Alberta
Edmonton, Alberta, Canada

Abstract: In recent years, the view that Information and Communication Technology (ICT) is vital in K-12 education has become widespread. ICT use in schools has increased and various professional bodies have set ICT standards for students and teachers. Schools of education are under pressure to produce teachers who are able to effectively integrate technology into their teaching. However, most teacher preparation programs do not adequately prepare teachers in ICT, nor assess candidates relative to ICT standards. This paper discusses the development of a computerized system to assess ICT declarative and procedural knowledge and to provide a profile to the participant.

Introduction

In recent years, governments, education organizations, and researchers have increasingly supported the view that incorporating ICT into learning and teaching is an important aspect of keeping the curriculum relevant and preparing students for their future in a complex knowledge-based world (Alberta Education, 1999b; CEO Forum on Education and Technology, 1997; Jonassen, 1995; Logan, 1995; Milken Exchange on Educational Technology, 1999; Thornburg, 1991). Data that provide insight into the computer literacy level of incoming undergraduate education students would be helpful to faculty designing appropriate curriculum. However, there is currently no assessment tool that establishes the ICT literacy level of current or prospective students. The predictive validity of a high school transcript or grade 12 English or mathematics marks is insufficient. How do the ICT skills of recent undergraduate students compare with the skills indicated at the Grade 12 level of the Provincial technology outcomes? Although the Best Practices in Technology documents (Alberta Education, 1999a) indicate that much ICT-related activity is occurring in schools, and an optional ICT K-12 curriculum has been available since June 1998, there are currently a few students entering educational technology courses at this university who have literally never turned on a computer (based on an informal "hands-up" survey of students in September 1999), while some students are familiar with some aspects of ICT, and a few are quite adept. This paper, part of a larger study (Davies, 2002) discusses the development of several Web-based instruments (Research Consent Form, Attitude Survey, Background Survey, Knowledge and Performance Tests) which assess the computer literacy level of incoming undergraduate education students.

Development of the Instruments

The Background Survey and Knowledge Test were implemented using ASPs written in the VBScript (Microsoft Corporation, 2000b) programming language. The ASP dynamically generated a Web page containing an HTML form with question and response data stored in a Microsoft Access 2000 relational database on a Windows server. The Background Survey included a variety of form field types that enabled implementation of different question types: radio button (multiple-choice/single response), drop-down list (multiple-choice/single response with a large set of possible answers), check box (multiple-choice/multiple response), and text box (short answer). The Knowledge Test was composed entirely of multiple-choice questions with five possible radio button answers.

The ASP also inserted client-side JavaScript (Netscape Communications Corporation, 2000) code into these HTML pages for the purpose of quick data validation. When the participant clicked on the form "Submit" button, the JavaScript local editing procedure was invoked. If unacceptable data were found, an error message window was displayed on the screen and the form was not submitted. For example, the Background Survey asked for the year of high school graduation. Acceptable responses had to be a four-digit year not greater than the current year. Upon acknowledgment of the error message, the participants' display was automatically scrolled to the question where the error was found. The JavaScript procedure also checked to see whether all applicable questions had been answered. If missing responses were found, a warning message window was displayed on the screen. Since it was unethical to demand that the participant answer all questions, the participant was then given the choice to return to the first missing response or to submit the data. The Knowledge Test responses were immediately scored by comparing with the correct response stored in the database for each item.

Observations of participants filling in the Web-based forms indicated that the JavaScript validation routines contributed to the completeness and accuracy of the data collected. For example, one participant was observed to receive the warning message stating that not all questions were answered. The individual read the warning message, returned to the form, stated "Oops, I missed that question," clicked on a response, and proceeded to submit the form with all questions completed. In all of the Background Survey data collected during the pre-course pilot test (34 participants x 35 questions each for a total of 1190 items) only 1 missing response and no invalid responses occurred. It would have been extremely time-consuming to ensure this level of data completeness and accuracy with paper-based forms. In the corresponding Knowledge Test data (952 items), there were 45 unanswered items. These were treated as incorrect responses when computing the overall test score. The vast majority of the missing responses (39) occurred because two participants attempted the first few questions and then submitted the form without checking responses for the remaining questions. Some additional messages were added to the Knowledge Test to ensure that students would be aware that there was no penalty for guessing.

The Performance Test was a much more technically complex instrument than the Knowledge Test since it required automating the analysis of files that participants manipulated on their local computers. The actual applications tested were established based upon required ICT skills, but were delimited by criteria such as time constraints for the initial instrument development, minimizing problems in collecting data, and allowing students to take the Performance Test with varying versions of application software on either the Windows or Macintosh platform. The solution chosen was to create a Visual Basic for Applications (VBA) (Microsoft Corporation, 2000) procedure within the same Access database described earlier. Web-based (especially client-side) programming techniques were avoided because of variable client computer setup and security issues involved in attempting to examine files on a client computer over the Internet. This part of the system required exchanging a set of files (compressed into a single archive) between each participant's computer and the database server.

During the Performance Test, the VBA procedure was continually running, monitoring a certain file directory every 60 seconds for arriving submissions and executing an automated scoring routine. The VBA scoring procedure implemented programming techniques (e.g., use of Microsoft Automation objects, methods and properties) which enabled automated execution of file system commands (e.g., file searches and directory listings), reading of text stream files, interfacing with external applications (e.g., Microsoft Word and Excel), opening files in these programs, and examining their object hierarchy.

The automated scoring routines were subjected to several iterations of testing and refining. The first author created about a dozen test cases - sets of computer files representing completed student practical tests - with a mixture of completely correct, partially correct and completely incorrect solutions for each task, completed using different software versions on different computer platforms. These were subjected to the automated routines and the scores for each item of each test case were manually checked for accuracy. This was repeated until all of the test cases were being properly scored. These same processes were repeated with files created by various expert reviewers as well as the entire group of student files resulting from the performance pilot test. Programming efficiency was also examined and improved until the time required to score the performance test averaged less than a second per test case.

The procedure that scored the spreadsheet activity of the Performance Test highlights the flexibility in the software used by participants allowed by the programming solution chosen. Spreadsheet files created in Excel 5 and 98 for Macintosh and Excel 95, 97, and 2000 for Windows formats were all scored without technical problems. In addition, the procedure also worked for files created in other programs such as Apple/ClarisWorks or Corel Quattro Pro then saved in Excel format. Multi-format flexibility assumes that the spreadsheet activities chosen for the test are limited to common features available in the different spreadsheet file formats. Only common tasks such as basic text or number formatting, cell alignment, and formulas were used. This was adequate for the level of expertise being measured in the target population.

Pilot Test

The instruments were pilot-tested on a thirty-five volunteer undergraduate education students prior to the main data-gathering period. These students were drawn from registrants in two Summer 2000 sections of the recommended educational computing option course. The students completed the Web-based forms and the Performance Test in a campus lab on Pentium 450 MHz computers with Microsoft Windows 98 and Office 2000 (Microsoft Corporation, 2000a) installed. No student chose the options of using a Macintosh computer, or to fill in paper-based copies of the online forms. The same group of students was later invited to participate in a Knowledge and Performance post-test, which was held on the second-last day of the term. After completing the test, the students were given a list of answers to the multiple-choice Knowledge Test, computer files providing correct solutions to the Performance Test, and personal help with any questions they had.

Instrument Validity

Content validity of the instruments was independently judged by three individuals who have expertise in educational technology: a faculty member, a PhD student, and a senior undergraduate student who had worked for a year as a marker in the educational technology undergraduate course. A number of modifications were made to the instruments based on the feedback from these initial reviewers. For example, some items that were deemed inappropriate were deleted from the instruments, the wording of some items was clarified, computer displays were improved, and new items suggested by the reviewers were added. These reviewers also served to verify that the system was operating without technical errors. Additional educational technology experts were called upon to similarly review the instruments after the pre-course pilot test. Improvements to the instruments were made as a result of this second round of validation activities.

Feedback on the instruments was obtained from the pilot-test students in a number of ways. First, during the pre-test, the researcher asked the students to raise their hand if at any time during the testing they found any information on the consent form or any question on the instruments to be unclear or inappropriate. A few such questions occurred and were discussed privately with the participant. These inquiries were noted on the researcher's printed copies of the instruments. Second, after reviewing the pre-test data, several students were contacted by email and asked for more information concerning their answers to certain items on the Background Survey. This resulted in some ideas for additional changes to the survey. Third, during the post-test, the students were asked to fill in a short feedback sheet. They were asked whether there were items that they felt were unclear or inappropriate on either the Knowledge or Performance Test, and whether they had any suggestions for additional items that could be included. All of these sources of student feedback were reviewed and resulted in modifications to the instruments.

Statistical correlations were computed as indicators of test validity (see Table 1). A high correlation ($r(33) = .460$, $p<.01$)[1] between the Knowledge pre-test and Performance pre-test scores provided evidence of concurrent criterion-related validity. That is, there was logically some commonality in the underlying constructs that these two tests measured. Strong correlations between the course midterm exam which occurred two weeks after pre-test and the pre-test Knowledge ($r(30) = .533, p<.01$) and Performance ($r(30) = .606, p<.001$) scores were evidence of predictive criterion-related validity.

Correlations between the course final exam and the post-tests (run 1 day before the exam) were: Knowledge post-test ($r(21) = .409$, $p = .059$) and Performance post-test ($r(23) = .673$, $p<.001$). The latter correlation was significant and offers strong evidence of concurrent criterion-related validity. The first correlation, while not quite statistically significant at the .05 level ($p=.059$), still offered some evidence of validity. It should be noted that the course exams during the Summer 2000 term were entirely performance-based, thus it was not surprising that the correlations between the course exams were stronger with the Performance Test than with the Knowledge Test. Also, it was easier to obtain some marks by sheer guessing on a multiple-choice test than it was on a performance-based test. Correlations between the course final exam and the pre-tests were: Knowledge Test ($r(28) = .522$, $p<.01$) and Performance Test ($r(28)= .688$, $p<.001$). All of these correlations were further evidence of predictive criterion-related validity.

	Know Pre	**Perf Pre**	**Know Post**	**Perf Post**	**Course MExam**	**Course FExam**	**Course Assign**	**Course Total**
Know Pre	1.000							
Perf Pre	.460**	1.000						
Know Post	.556**	.332	1.000					
Perf Post	.342	.678**	.220	1.000				
Course MExam	.533**	.606**	.211	.162	1.000			
Course FExam	.522**	.688**	.409	.673**	.627**	1.000		
Course Assign	.506**	.486**	.141	.427*	.579**	.592**	1.000	
Course Total	.592**	.707**	.341	.543**	.804**	.917**	.808**	1.000

** Correlation significant at the .01 level * Correlation significant at the .05 level

Table 1: Pilot Test Instrument Validity - Correlation Statistics

Instrument Reliability

The Background Survey was not constructed as a scale where all items contributed to an overall score. Rather than computing internal consistency statistics, reliability of this instrument was established by selective re-testing. Four of the students who had volunteered to participate in the initial instrument evaluation were re-tested using a different format for presenting the questions (as an interview rather than online written questions). No differences in the responses from the two forms were found, indicating high reliability (Fraenkel & Wallen, 1996) although these participants did offer a few suggestions for clarifying the wording of a few items. The consistency of responses was not surprising, since most of the questions on the survey would be considered objective (mainly factual information such as whether or not they own a home computer). Answers to questions like these are likely to be answered the same in a test-retest situation where there is little time between tests.

Cronbach's alpha coefficient was computed for the pre-test Attitude Survey (.63), the Knowledge Test (.79), and the Performance Test (.91). The Attitude Survey reliability was judged too low for meaningful data interpretation (a widely accepted lower limit for alpha is .7), the second was acceptable, and the latter was exceptionally high, being at the level of marketed achievement tests (Fraenkel & Wallen, 1996, p.163).

[1] In this notation, r(33) means r(df), where df is the degrees of freedom (equal to n-1)

Improvements to the Attitude Survey were essential to establish solid reliability. Analysis of the inter-item correlation matrix identified three items that were poorly associated with the other items and thus did not contribute well to the overall test score; these items were modified. In addition, since the survey originally consisted of only 12 items, reliability could be easily raised by increasing the number of related items (Fraenkel & Wallen, 1996, p.163). A target of 20 items was established.

The alpha coefficients for the post-test Knowledge Test and Performance Test (run again at the end of term as a post-test) were not as high, .63 and .84 respectively. This was because the tests included some questions deemed by experts to be easy (equivalent to the stated prerequisites for the recommended educational computing course), yet which stumped many students in the pre-test, effectively screening individuals with very low knowledge or practical skills. Tests are not always equally effective in different situations (Murphy & Davidshofer, 1991); these tests were less effective as a post-test after completing a course which covers much of the content of the tests and provides remediation for missing prerequisite skills.

In the pre-course Knowledge Test, no questions (out of 28) were answered correctly by all students. In fact, the easiest question was answered correctly by 88% of students. By contrast, in the post Knowledge Test, there were 4 questions answered correctly by all students, and another 8 questions answered correctly by at least 75% of students. There was less overall variance in the post-test scores than in the pre-test scores. Also, items of zero variance (same score for all students) cannot be correlated with other test items and thus do not enter into the reliability calculations, which reduces the reliability coefficient (a measure of average inter-item correlation).

The course appears to have been effective in raising the student scores. The mean pre-to-post gain on the Knowledge Test was 17.52. A one-sample t-test on the gain scores (post – pre), comparing them to a test value of zero (equivalent to a dependent or paired samples t-test using the pre and post scores) found the difference statistically significant. On the Performance pre-test, no questions (out of 24) were answered correctly by all students, while on the post-test, 6 questions were answered correctly by all students. The mean pre-to-post gain on the Performance Test was 26.07; the one-sample t-test on the gain scores compared to zero found the difference statistically significant. It should, however, be noted that the difference in cases between the pre and post-tests must be considered; the students who didn't participate in both tests were not part of the gain analysis and may have differed from those that did.

The pre-to-post gains on these two tests affirmed the effectiveness of the instruction and were additional pieces of evidence for the validity of the tests. It demonstrated that participants who have had more training or practice with ICT tools (i.e., the students at post-test time) scored much higher than those with less (i.e., the students at pre-test time). This is logically consistent with what the tests purport to measure.

Time Required to Administer Instruments

The pilot pre-test also served to verify that the instruments could be completed within a reasonable timeframe on a single day. The actual time required for students to complete all of the forms and tests was approximately 1.5 hours, about 0.5 hour for all of the online forms (consent, attitudes, and knowledge) and 1 hour for the performance tasks. The start and end time for each participant's work on each online form was stored in the database, making it simple to calculate the average time required to complete a form. In the case of the Background Survey, the average time was about 8 minutes, and the maximum time was 12 minutes. The maximum time required for the 12 Attitude items was around 4 minutes. For the Knowledge Test, the minimum time required was 3 minutes, the maximum 25 minutes, and the average 13 minutes. For the Practical test the minimum time required was 12 minutes, the maximum 65 minutes, and the average 36 minutes.

Observations from the Pilot Test

A number of statistics were computed to provide a general picture of the pilot data. Histograms of the pre and post Knowledge and Performance test scores revealed that these distributions were approximately normal, with the post-tests having much higher means and lower variances than the corresponding pre-test. Overall, the achievement on the pre-course tests was quite low with the means of both tests being below 50%. Comparing the test questions against the recommended educational computing course curriculum and stated prerequisites, the researchers concluded that marginally acceptable course prerequisite skills and knowledge (basic computer operation, file management and word processing) would be indicated by a score of at least 50% on the pre-tests. On the Knowledge Test, 44% of students did not meet this standard and 59% did not meet it on the Performance Test, indicating that many students do not possess adequate course prerequisites. A more comfortable level of prerequisite skills would be indicated by scores of at least 60%. On the Knowledge Test, 77% of students did not reach the 60% level and 71% did not meet it on the Performance Test. At the other end of the spectrum, a few students performed well enough on the pre-test to be likely candidates for successfully challenging the course or taking courses that require skills equivalent to completing that course as a prerequisite. Obviously more testing and validity evaluation would be required to establish a standard for this, but in the authors' opinion, 80% seems like a level that would reasonably indicate mastery. If this were the case, 3% of students (1 individual in the pilot group) would have qualified. A summary report on the pilot test group performance on the pre-course tests was provided for informational purposes to the course instructors and senior teaching assistants.

Full Implementation of the Tests

The instruments were successfully incorporated into the regular offering of the course beginning in the Fall 2000 (approximately 1000 students). Validity and reliability tests gave similar results to the pilot test. In addition, the implementation of online testing substantially reduced the amount of time required to evaluate students, particularly in the Performance test. The data from these tests have established a baseline of data on the ICT skills and knowledge of students entering our Faculty of Education. We expect the average ICT skills and knowledge of education entrants to rise over the next few years. At some point a certain level of ICT Literacy could be an admission requirement. These automated instruments could be used as an efficient admission screening tool. The teacher education program could then focus more resources on improving teacher candidates' abilities to integrate technology into their teaching instead of development of basic ICT skills and knowledge.

REFERENCES

Alberta Education. (1999a). Best practices in technology. Retrieved March 27, 2000 from the World Wide Web: http://ednet.edc.gov.ab.ca/technology/bestpractices/bestpractices99.asp.

Alberta Education. (1999b). Learning technology in Alberta's schools: Information for parents. Retrieved March 27, 2000 from the World Wide Web: http://ednet.edc.gov.ab.ca/techoutcomes/.

CEO Forum on Education and Technology. (1997). From pillars to progress (School technology and readiness report - year one). Retrieved March 27, 2000 from the World Wide Web: http://www.ceoforum.org/downloads/97report.pdf.

Davies, J.E. (2002). Assessing and Predicting Information and Communication Technology Literacy in Education Undergraduates. Unpublished Doctoral Dissertation. Edmonton, Alberta: University of Alberta.

Fraenkel, J. R., & Wallen, N. E. (1996). *How to design and evaluate research in education* (3rd ed.). New York: McGraw-Hill.

Jonassen, D. H. (1995). Supporting communities of learners with technology: A vision for integrating technology with learning in schools. *Educational Technology, 35*(4), 60-63.

Logan, R. K. (1995). *The fifth language: Learning a living in the computer age*. Toronto, On: Stoddart.

Microsoft Corporation. (2000). Microsoft Office 2000 [Productivity software suite]. Redmond, WA.

Milken Exchange on Educational Technology. (1999). Transforming learning through technology: Policy roadmaps for the nation's governors. Retrieved March 27, 2000 from the World Wide Web: http://www.milkenexchange.org/project/nga/ME266.pdf.

Murphy, K. R., & Davidshofer, C. O. (1991). *Psychological testing: Principles & applications* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Netscape Communications Corporation. (2000). JavaScript [Computer programming language]. Mountain View, CA.

Thornburg, D. D. (1991). *Education, technology, and paradigms of change for the 21st century.* Eugene, OR: Starsong Publications.