DOCUMENT RESUME

ED 366 507 SE 054 134

AUTHOR Helgeson, Stanley L.; Kumar, David D.
TITLE A Review of Educational Technology in Science Assessment. Monograph Series Number 7.
INSTITUTION National Center for Science Teaching and Learning, Columbus, OH.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE Nov 93
CONTRACT R117Q00062
NOTE 11p.
AVAILABLE FROM National Center for Science Teaching and Learning, 1929 Kenny Road, Columbus, OH 43210-1080.
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Computer Assisted Testing; Computer Uses in Education; *Educational Technology; Elementary Secondary Education; Higher Education; Homework; *Hypermedia; Science Activities; Science Education; Science Experiments; Science Tests; *Student Evaluation
IDENTIFIERS Authentic Assessment

ABSTRACT

Emerging applications of microcomputers and hypermedia to assessment in science education are reviewed. Although the current use of technology consists mainly of computerized administration of multiple choice tests drawn from item banks, the potential advantages are much greater. Among these advantages are immediate feedback to students, formative evaluation with remediation possibilities, adaptive testing in which the test is adjusted to match the students' level of performance, monitoring of homework, and laboratory activities. (Contains 42 references.) (Author/AA)

A Review of Educational Technology in Science Assessment

Stanley L. Helgeson & David D. Kumar

THE NATIONAL CENTER FOR SCIENCE TEACHING & LEARNING

1929 Kenny Road
Columbus, OH 43210
(614) 292-3339

A Review of Educational Technology in Science Assessment
Supported by Grant #R117Q00062 from the Office of Educational Research and Improvements, United States Department of Education
A Publication of the National Center for Science Teaching and Learning
Monograph Series Number 7, November 1993

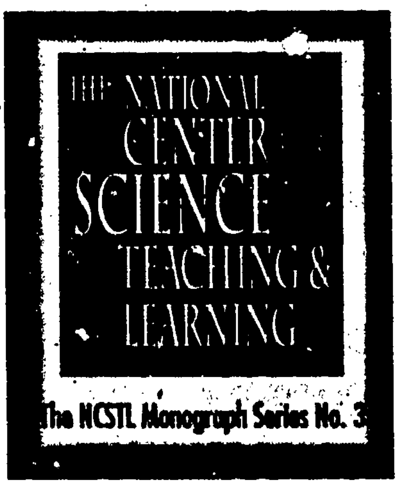
A Review of Educational Technology in Science Assessment

Stanley L. Helgeson & David D. Kumar

Emerging applications of microcomputers and hypermedia to assessment in science education are reviewed. Although the current use of technology consists mainly of computerized administration of multiple choice tests drawn from item banks, the potential advantages are much greater. Immediate feedback to students, formative evaluation with remediation possibilities, adaptive testing in which the test is adjusted to match the student's level of performance, monitoring of homework, and laboratory activities.

Introduction

The purpose of this review was to examine and summarize the research literature pertaining to the role of educational technology in science education assessment. Educational technology has been a center of development and research in science teaching and learning (Grandgenett, Ziebarth, Koneck, Farnham, McQuillan, & Larson, 1992; Kumar, 1991a). Similarly, the search for alternative assessment strategies has been a focus of activities and developments in educational testing and evaluation (Stiggins and Bridgeford, 1985; Shavelson, Carey, & Webb, 1990; Swain, 1991; King & Bathwaite, 1991). "According to some scientists, the true test of students' understanding is to put them in a laboratory, pose a problem, and let them use the resources of the lab to solve the problem" (Shavelson, Carey, & Webb, 1990, p. 696). However, "large-scale hands-on testing in laboratories is far too costly in time, dollars, human resources, and equipment" (p. 696). Therefore, according to Shavelson, Carey and Webb (1990) "Researchers, in partnership with practitioners, need to build a new knowledge base and a new technology for achievement testing in science" (p. 693).



When considering a new knowledge base for the assessment of the processes of learning and problem solving in the light of educational technology one cannot overlook the role of the developments in cognitive psychology. There appears to be a strong relationship between the developments in computer technology, and cognitive psychology (De May, 1992). Cognitive theories in combination with educational technology, especially the hypermedia, offer promises to meet the challenges of the assessment reform calls in science education. Techniques such as concept mapping and cognitive task analysis have a profound role to play in analyzing learning processes, and they provide a means to understand the structure of human knowledge with the assistance of educational technology (Bower & Hilgard, 1981; De May, 1992). For example, one of the ways of representing semantic knowledge using computers involves the use of "nodes" to represent concepts in terms of texts and graphics and "links" to represent the semantic relationships between the nodes which is the underlying framework of the hypermedia technology (Halasz, 1988). Thus an argument could be made that, since the hypermedia can also be used to represent human knowledge structure, it can also be used as a medium for understanding human cognitive processes (Kumar, 1992) [e.g., moves and decisions in a problem space (Collins, 1990)].

Educational technologies such as computers and hypermedia are in the forefront, and they are "the closest approximation to hands-on performance evaluation that can be group administered" (Shavelson, et al, 1990, p. 5). For example, computers and hypermedia applications could provide multi-dimensional environments to study the process of learning and problem solving, and to represent knowledge structures (Jonassen, 1988; Champagne, & Klopfer, 1984; Bower & Hilgard, 1981). Thus, computers and hypermedia not only find applications in the development of alternative assessment technologies but also provide environments for understanding the processes involved in assessment in science education.

Method of Document Selection

Documents were initially identified for this review by conducting a search of the ERIC data base. Search terms used were all possible permutations of computers, assessment, testing, hypermedia, and science education.

Next, documents were identified from known sources, including the references from these documents. Each document thus identified was then subjected to a systematic review and those articles dealing with computer and hypermedia applications to assessment in science education were selected for inclusion.

Computer Applications in Assessment

In summarizing research findings on computer-based education, Waugh and Currier (1986) found that: (1) groups experiencing some kind of computer-based education attained test scores which were on average between .25 and .44 standard deviations higher than their comparison groups; (2) there was evidence favoring the use of computer-based education with academically disadvantaged students; (3) long term retention was no better for computer-based education than for other modes of instruction; (4) secondary students who experienced computer-based education had more positive attitudes toward computers than did their peers who did not experience computer-based education; and, (5) there was significantly less time required for computer-based education compared to conventional instruction. It should be noted that many of the studies summarized relied heavily on drill and practice modes of instruction. Such programs depend upon immediate feedback as a major function. While this may not fit the common perception of assessment, it appears that it does in fact function in such a manner and that the immediate feedback may well have a positive impact on learning.

A common use of computers in assessment is to provide teachers with access to large banks of items for testing. These may range from specific topics such as medical biochemistry (Aesche & Parslow, 1988) for instructors of a given course, to a test bank designed for state assessment (Willis, 1988), to a broad range of juried test items which teachers anywhere in the country may access and download into their own computers (Dawson, 1987). Once the item banks are in place, the computer may then be used to devise unique combinations of test items for each student and to use the results of those tests to develop remedial learning activities for each student. In each case, the computer can administer the quizzes, grade and record the results, and provide the student with immediate feedback (Dunkleberger, 1980). Use of the computer to file test questions, assemble examinations, handle all records, produce and grade tests and guide students to what should be done next enables testing to be done with an efficiency not possible from any teacher (Summers, 1984; Vogel, 1985; Heikkinen & Dunkleberger, 1985). Specially designed punch cards can be used for testing and grading large populations. The cards can be processed routinely at batch process stations by lab personnel who have little computer knowledge. This allows for cheap and easy marking and can be adapted to a wide variety of tests (Mihkelson et al., 1984). The availability of microcomputers and test item banks makes possible the transition from punch cards to computer-based assessment with all the advantages indicated for punch cards.

A form of formative assessment makes use of the computer to evaluate student data collected in laboratory exercises. Such checking of data and calculations is repetitive, prone to error, and not cost effective when done by humans. Computers, on the other hand, excel at this type of task (Harrison & Pitre, 1983, and Harrison & Pitre, 1988). Programs used in this way are designed to check for realistic values, a range of data, and values clearly outside acceptable limits. When incorrect answers are given, students may be asked to redo their calculations and submit revised figures (May, Murray, & Williams, 1985). The programs also may be designed to tentatively accept answers within a certain range, but to suggest that students return to places of potential error and check their work (Harrison & Pitre, 1988).

A common use of computers in assessment is to provide teachers with access to large banks of items for testing.

As part of a project to integrate computer-generated homework into physical science college courses, Milkent and Roth (1989) used computer-generated problems as homework assignments and monitored student progress with computer-generated multiple choice quizzes. They found that the use of the computer-generated homework significantly reduced the effectiveness of ACT scores as predictors of course achievement. Put in other words, as a result of the homework approach, students had greater opportunities for achieving mastery and for minimizing the potential influence of entry level aptitude and prior academic preparation. This was in addition to the teacher advantages of an efficient system for homework management and freedom from bookkeeping procedures.

Incorporation of computers into science instruction often takes the form of microcomputer-based laboratories (MBL). Assessment is frequently a part of such a system. However, in some cases this means simply presenting multiple choice questions by means of the computer screen (Bross, 1986). If immediate feedback is not available, no learning gains may accrue to such computer use. Increased ease of data collection and processing may still make this approach to testing of value to the instructor. A more useful approach might be that described by

Browning and Lehman (1988) for identifying student misconceptions in genetics problem solving. Four computer programs were presented and the students' responses were recorded and analyzed for evidence of misconceptions and difficulties in the problem solving process. Three main problem areas were identified: difficulties with computational skills, difficulties in the determination of gametes, and inappropriate application of previous learning to new problems. Evaluation of this type would seem to show considerable promise for remedial instruction and improved student learning.

The students indicated that most of them favored the computer-administered tests . . .

Collins (1984) conducted a study to determine whether learning would be improved with computerized tests. Two hundred ten students were enrolled in a one-semester introductory biology course. Students in the computer section took computer generated tests in addition to the tests taken by students in the other sections. Students taking the computer tests were given immediate feedback on their scores, then told which responses were correct and which were incorrect. In addition, the computer recorded student data on disk, allowing for later analysis by the instructor. Collins concluded that computer testing led to enhanced learning as indicated by higher scores on weekly in-class written tests, the midterm examination, the final examination, and final class marks.

Collins and Earle (1989–90) examined the effects of computer-based learning and computer-administered testing in an introductory biology class. They found that the greatest benefit was attained by those using the computer units in addition to attending regular lectures. Taking weekly computer-administered multiple choice tests also appeared to benefit students of middle and upper ability but not students of lower ability levels. That the use of weekly computer-tests can increase students' scores reinforces a finding of an earlier study (Collins, 1984). Although students benefited from using either the computer learning units or the computer tests, the use of the two together did not result in even more gain, as might have been expected. Frequency of use of the units appeared to be a factor in that the "frequent" user group achieved a much higher mean score and higher pass rate than did the "infrequent" user group.

Such checking of data and calculations is repetitive, prone to error, and not cost effective when done by humans.

The possibility that students were being disadvantaged by taking computer tests instead of written paper forms of the same tests was studied by Fletcher and Collins (1986–87). They found that students' mean scores on the computer-administered test and the written forms of the same test were roughly equivalent, and concluded that the students were not disadvantaged by taking the computer tests. The students indicated that most of them favored the computer-administered tests and cited several major advantages: (1) immediacy of scoring; (2) immediate feedback on incorrect answers; (3) more convenient, straight forward and easy-to-use; and (4) faster than written tests. Two major disadvantages were noted by the students: (1) not being to review all their responses at the end of the test and make changes and, (2) not being able to skip questions and come back to answer them later (p. 42).

The converse case was studied by Jackson (1988) who attempted to discover whether a computer could give any significant educational advantage to the pupil. That is, could the computer improve pupil motivation during the test, by giving instant feedback and marking, thus improving understanding and hence give an enhanced score in a future test? (p. 809) The middle school science students who were tested by computer and given immediate feedback scored significantly higher in a later test using the same material than did those students who were tested using the traditional paper and pencil method. An additional gain for the teacher was the ability to conduct further analyses, such as test item analysis, on the computer-recorded student data; such analyses could not be easily carried out without computer administered testing.

Computerized adaptive testing is emerging as a more efficient way to assess student knowledge. A unique characteristic of this technique is that each examinee is given an individualized test comprised of questions from a content-valid item bank. The adaptive algorithm selects questions that provide the most information about the examinee given his/her current estimated ability measure. After answering each question, the examinee's ability is re-estimated. If the correct answer is given, the examinee's measure increases and the next question is more difficult. If an incorrect response is submitted, the measure decreases and the next item administered is easier. This results in a test that is tailored to each individual. The tests can be of various lengths depending upon how far above or below the pass/fail threshold the examinee's performance falls. Thus, a test sufficiently long to clearly determine the best decision can be presented with no wasted questions. A pilot study of the effectiveness of computerized adaptive testing for certification in five medical technology fields revealed that

. . . the use of microcomputer-administered diagnostic testing was successful in increasing student achievement . . .

50 to 100 questions served to provide the necessary pass/fail information as compared to 109 written questions. The computerized test took two to two and a half hours to complete compared to four hours for the written test. Other benefits of computerized testing included a shorter turn-around time of test results, improved security and data collection, and less chance of cheating due to the individualized nature of the exams (Herb, 1992).

The effects of microcomputer-administered diagnostic testing on both student achievement and attitudes were of concern to Waugh (1985). Students in one group were given the unit objectives and responded to a computer-administered diagnostic test consisting of one item per objective. The other group received the objectives and were assigned an out-of-class task of completing an objective-specific mini-project. The results showed that microcomputer-administered diagnostic testing could positively influence the immediate achievement of students in science. Evidence did not, however, support the hypothesis that an exposure to diagnostic testing might influence continuing achievement. The findings indicated that the use of microcomputer-administered diagnostic testing was successful in increasing student achievement in science by an average of six percent with no loss of positive attitude toward school, learning, or science. The evidence further indicated that diagnostic testing might have played a role in arousing student interest in microcomputers.

Student attitudes were also the focus of a study by Knight and Dunkleberger (1977) in a comparison of computer-managed self-paced instruction with teacher-managed group-paced instruction for ninth grade students. The course consisted of large group lectures (31 percent of the overall time), small group seminars (46 percent of the time), and laboratory activities (31 percent of the time). The computer-managed self-paced group and the teacher-managed group-paced students received the same large group lectures and small group seminars. The computer group was allowed to self-pace through the laboratory activities while the teacher-managed group followed a group-pace. The computer served as an assessment and record keeping device for the computer-managed students. The quizzes were four-choice, multiple choice questions and students received immediate feedback after completing each item. Although the differing instructional approaches were applied only during the laboratory component of the course (31 percent), the positive reaction of the computer-managed self-paced group was sufficiently strong to effect a significant difference in attitudes toward the study of science.

Hypermedia in Assessment

The impact of emerging interactive videodisc technology was studied by Huang and Aloi (1991) in a first year biology course. The interactive video involved 17 menu-driven chapters integrating computer text with laser disc images and computer graphics. The students were organized into groups with inter-group competition in answering true/false, multiple choice, and completion questions. The researchers compared, using an unpaired t-test, the proportion of students getting A, B, C, D, F, and W (withdraw) for 11 semesters prior to using interactive video with the proportions during the five semesters following its use. They found that the proportion receiving A's increased significantly ($p<.005$) following use of the interactive video. The percentage increases were: A's, 6 percent before and 18 percent after; B's, 21 percent before and 32 percent after; C's, 20 percent before and 36 percent after; D's 10 percent before and 4 percent after; F's did not change. Retention of students was also increased. The

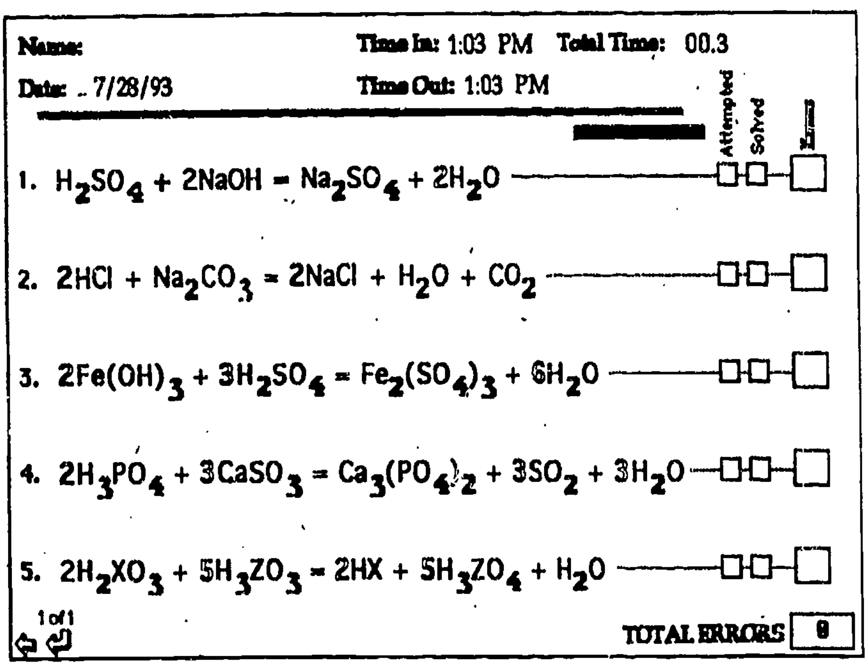


Figure 1

proportion of withdrawals was 33 percent before interactive video use and 24 percent after. Thus, the use of interactive videodisc resulted in increased proportions of success at nearly all levels of achievement.

Interactive videodisc (IVD) was also used as a tool in assessing science teachers' knowledge of safety regulations in school laboratories for purposes of teacher certification by the Connecticut State Department of Education (Lomask, Jacobson, & Hafner, 1992). The program simulates a typical lab activity in a secondary school general science course and shows four students performing a simple lab experiment to identify unknown materials. The IVD assessment includes two stages: stage one deals with safety equipment and storage of chemicals and stage two deals with students' laboratory practices. The examinees are asked to assume the role of the lab teacher by viewing an interactive videodisc simulated classroom. The teachers are then asked to identify safety violations and to suggest preventive or corrective measures. Subjects' responses are recorded for later analysis and scoring (p. 1).

An emerging application of hypermedia in assessment involving problem-based learning in chemistry is found in the "Hyperequation" (Kumar, 1991b) project at the National Center for Science Teaching and Learning at The Ohio State University. Hyperequation is an assessment software developed in HyperCard™ on a Macintosh platform to study student performance in balancing stoichiometric chemical equations (see Figure 1).

Hyperequation (in its pilot stage) has the following features. It is easy to operate through the computer-mouse interface. It has been programmed to provide immediate feedback and motivation, and to register some pertinent information involved in the process of balancing stoichiometric equations. One of the purposes of this software is to simulate similar tasks involving traditional paper-pencil methods of assessment, in addition to providing a non-linear visual environment for problem solving. For example, the Hyperequation can keep a record of the number of attempts and the order with which responses are made by each student including the total time on-task. Also Hyperequation can display on screen as well as provide a printout of an overall and item-by-item record of each student's performance on the problem task (see Figures 2 and 3). Due to confidentiality of student performance records, only the classroom teacher, through a password, has access to this information in Hyperequation.

Prima facie evidence from a pilot study involving the task of balancing chemical equations using the

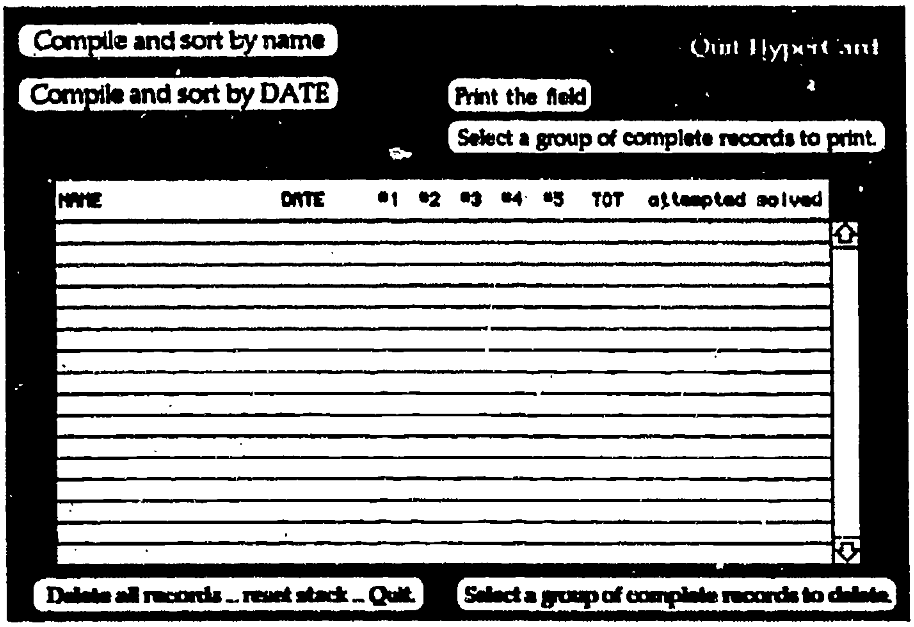


Figure 2

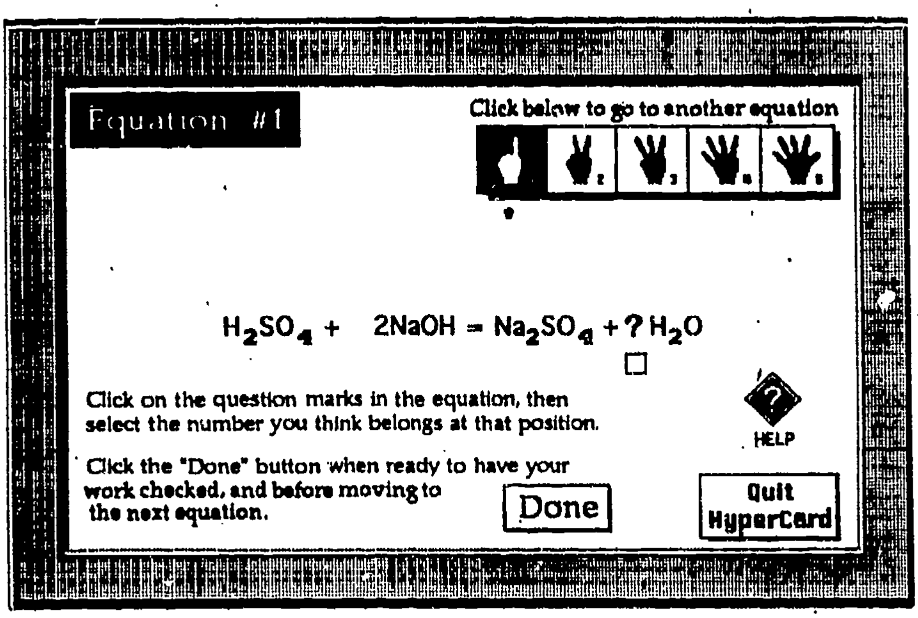


Figure 3

Another form of formative assessment is made possible through the use of computers to monitor homework and laboratory activities.

HyperCard™ method (Hyperequation) described above and traditional pen-paper assessment methods indicates that the HyperCard™ method influenced differently the performance of expert and novice students in balancing stoichiometric chemistry equations (Kumar, White, & Helgeson, 1993). Possibly the presence of a computer including the flexible environment of Hyperequation may be the reason for the difference. Maybe the hypermedia environment of Hyperequation helps novices to perform better with HyperCard™ than with the traditional method due to the following reasons: The mouse-interface with the computer was perhaps less interfering than the pen-interface with the paper in solving the stoichiometric chemical equations; the use of HyperCard™ may tend to reduce the initial differences in student expertise (Milkent and Roth, 1989); the computer itself provided an added external memory for the student while balancing the equations thereby reducing the cognitive demand on working memory; and the HyperCard™ method provided immediate feedback so that the student was motivated to stay on task until a satisfactory solution was reached.

Martinez (1991) has reported a similar hypermedia environment using an "IBM-compatible computer interface delivery" platform for administering "figural response" test items to cell and molecular biology students. With a computer-mouse interface, a set of computer screen tools are activated by buttons (e.g., "move object," "rotate," "draw line"). For example, chromosomes and molecular groups are moved on the screen by students to respond to various questions such as "Given the D-glucose below, construct its L-glucose stereoisomer using the template shown" (p. 387).

A similar work in physics at the University of California-Santa Barbara in collaboration with the California Institute of Technology has been reported by Shavelson, Baxter, Pine, Yure, Goldman, and Smith (1990). For example, using a simulation "Electric Mysteries" on a Macintosh platform, a hands-on environment for assessment in electric circuits was replicated. Students have to find out the circuitry among five possible circuit designs from five "mystery boxes" by manipulating icons on the Macintosh computer, instead of physically manipulating bulbs, batteries, and wires. Every move made by the student is recorded by the computer which is later used for assessment. The findings indicate that expert students performed significantly better on the electric mysteries problem than novices.

Summary

There appear to be several advantages to incorporating some form of computer assistance in assessment. Immediate feedback to the students seems to be a consistent factor in increased achievement. Ease of test taking, together with improved record keeping, suggest improved efficiency for both students and teachers. The availability of large test item banks makes possible several intermediate quizzes with achievement gains appearing to result from this practice. Another form of formative assessment is made possible through the use of computers to monitor homework and laboratory activities. Such formative evaluation serves both as a diagnostic tool and as a remediation device, indicating where corrections are needed. The data collection capability of computer testing also permits more extensive data analysis, especially in the area of test item analysis, which in turn should yield more reliable and, presumably, more valid assessment. Two cautions must be noted, however. First, the simplicity of devising multiple choice, true/false, matching, and other objective tests can lull the teacher into simply doing a better job of assessing low level recall knowledge. Second, the linear nature of most computer testing does not allow the student to go back and reflect upon a particular item, nor to view the completed test as a whole to check for consistency of responses. The increased improvement and implementation of such emerging technologies as interactive video and hypermedia (Kumar, 1991a) show high promise for overcoming both difficulties by providing opportunities for both improved levels of questions and increased flexibility in the testing process because of the non-linear capabilities inherent in hypermedia.

While the research evidence is still limited it appears that some tentative conclusions may be drawn. The first, and possibly most important, finding is the positive effect on achievement of immediate feedback and its attendant reinforcement. A second outcome is the increased ease and simplicity of test-taking and data collection and analysis. Next, there is an increased facility to do formative or intermediate assessment with accompanying remediation. Finally, with the emergence of hypermedia, there is increased flexibility of assessment allowing for a potentially better match between the way in which humans construct knowledge and methods for assessing such learning. However, as Linn, Baker and Dunbar (1991-1992) stated, more research is warranted to validate educational technology for performance assessment especially in issues related to gender and sociocultural factors, and the role of the classroom teacher in assessment in science education. More research and development in educational technology in science assessment can be expected to lead to novel applications and newer frontiers in science education.

Acknowledgements

Thanks to Robert Riordan for his assistance in a previous draft of this article.

References

Aesche, D. W., & Parslow, G. R. (1988). Use of questions from the medical biochemistry question bank with the 'Q' instruction package. *Biochemical Education, 16*(1), 24-28.

Bross, T. R. (1986). The microcomputer-based science laboratory. *Journal of Computers in Mathematics and Science Teaching, 5*(3), 16-18.

Browning, E., & Lehman, J. D. (1988). Identification of student misconceptions in genetics problem solving via computer program. *Journal of Research in Science Teaching, 25*(9), 747-761.

Bower, G. H., & Hilgard, E. R. (1981). *Theories of learning*. Englewood Cliffs, NJ: Prentice-Hall.

Champagne, A. B., & Klopfer, L. E. (1984). Research in science education: Cognitive psychology perspective. In D. Holdzkom & P. B. Lutz (Eds.), *Research within reach: Science Education*. Washington, DC: National Science Teachers Association.

Collins, M. A. J. (1984). Improved learning with computerized tests. *The American Biology Teacher, 46*(3), 188-190.

Collins, M. A. J., & Earle, P. (1989-90). A comparison of the effects of computer-administered testing, computer-based learning and lecture approaches on learning in an introductory biology course. *Journal of Computers in Mathematics and Science Teaching, 9*(2), 77-84.

Collins, A. (1990). Cognitive apprenticeship and instructional technology. In L. Idol & B. F. Jones (Eds), *Educational values and cognitive instruction: Implications for reform*. Hillsdale, NJ: Erlbaum.

Dawson, G. (1987). 10,000 science questions by modem. *The Science Teacher, 54*(3), 41-44.

De May, M. (1992). *The cognitive paradigm*. Chicago: The University of Chicago Press.

Dunkleberger, G. E. (1980). Applying 20/20 hindsight to self-pacing. *The Science Teacher. 47*(8), 32-34.

Fletcher, P., & Collins, M. A. J. (1986-87). Computer-administered versus written tests—advantages and disadvantages. *Journal of Computers in Mathematics and Science Teaching, 6*(2), 38-43.

Grandgenett, N., Ziebarth, R., Koneck, J., Farnham, M. L., McQuillan, J., & Larson, B. (1992). An investigation of the anticipated use of multimedia by pre-service teachers. *Journal of Educational Multimedia and Hypermedia, 1*, 91-101.

Halasz, F. G. (1988). Reflections on notecards: Seven issues for the next generation of hypermedia systems. *Communications of the ACM, 31*(7), 836-852.

Harrison, D., & Pitre, J. M. (1983, December). Computerized lab test on error analysis. *The Physics Teacher, 21*(9), 588-589.

Harrison, D., & Pitre, J. M. (1988). Computerized pre-checking of laboratory data. *The Physics Teacher, 26*(3), 156.

Heikkinen, H., & Dunkleberger, G. E. (1985). On disk with mastery learning. *The Science Teacher, 52*(7), 26-28.

Herb, A. (1992, February 17). Computer adaptive testing. *ADVANCE*, (Newsletter of the Medical Laboratory Professionals).

Huang, S. D., & Aloi, J. (1991). The impact of using interactive video in teaching general biology. *American Biology Teacher, 53*(5), 281-284.

Jackson, B. (1988). A comparison between computer-based and traditional assessment tests and their effects on pupil learning and scoring. *School Science Review, 69*(249), 809-815.

Jonassen, D. H. (1988). Designing structured hypertext and structuring access to hypertext. *Educational Technology, 28*(11), 13-16.

King, W., & Bathwaite, W. (1991). School-based assessment in science—a Caribbean perspective. *School Science Review, 72*(261), 127-131.

Knight, C. W., & Dunkleberger, G. E. (1977). The influence of computer-managed self-paced instruction on science attitudes of students. *Journal of Research in Science Teaching, 14*(6), 551-555.

Kumar, D. D. (1991a). Hypermedia: A tool for STS education? *Bulletin of Science, Technology & Society, 11*(6), 331-332.

Kumar, D. D. (1991b, September). *A Note on "Hyperequation,"* An occasional paper. National Center for Science Teaching and Learning, The Ohio State University, Columbus.

Kumar, D. D. (December 1992). *Can cognitive psychology shape educational technology for performance assessment?* An occasional paper. National Center for Science Teaching and Learning, The Ohio State University, Columbus.

Kumar, D. D., White, A. L., & Helgeson, S. L. (April 1993). *Effect of HyperCard and traditional performance assessment methods on expert-novice chemistry problem solving.* A paper accepted for presentation at the Annual Meeting of the National Association for Research in Science Teaching, Atlanta, Georgia.

Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991–1992). Complex, performance-based assessment: Expectations and validation criteria. *Evaluation Comment.* Los Angeles, CA: UCLA's Center for the Study of Evaluation & The National Center for Research on Evaluation, Standards and Student Testing.

Lomask, M., Jacobson, L. S., & Hafner, L. P. (1992, March 21–25). *Interactive Videodisc as a Tool for Assessing Science Teachers' Knowledge of Safety Regulations in School Laboratories.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Boston, MA.

Martinez, M. E. (1991). Figural response in science and technology testing. In G. Kulm & S. M. Malcom (Eds.), *Science assessment in the service of reform.* Washington, DC: AAAS.

May, P. M., Murray, K., & Williams, D. R. (1985). CAPE: a computer program to assist with practical assessment. *Journal of Chemical Education. 62*(4), 310-312.

Mihkelson, A., White, G. D., & Williams, A. J. (1984). "The Application of computer technology in the assessment of large undergraduate classes. *Journal of Chemical Education, 61*(7), 631.

Milkent, M. M., & Roth, W. M. (1989). Enhancing student achievement through computer-generated homework. *Journal of Research in Science Teaching, 26*(7), 567-573.

Shavelson, R. J., Carey, N. B., & Webb, N. M. (1990, May). Indicators of science achievement: Options for a powerful policy instrument. *Phi Delta Kappan*, 692-697.

Shavelson, R. J., Baxter, G. P., Pine, J., Yure, J., Goldman, S. R., & Smith, R. (1990, April). *Performance indicators for large-scale science assessment.* Paper presented at the annual meeting of the American Educational Research Association, Boston, MA.

Stiggins, R. J., & Bridgeford, N. J. (1985). The ecology of classroom assessment. *Journal of Educational Measurement, 22*(4), 271-286.

Summers, G. (1984). Testor. *Journal of College Science Teaching, 13*(5), 356-358.

Swain, J. R. L. (1991). The nature and assessment of scientific explorations in the classroom. *School Science Review, 72*(260), 65-77.

Vogel, G. (1985). A microcomputer-based system for filing test questions and assembling examinations. *Journal of Chemical Education, 62*(11), 1024-1027.

Waugh, M. L. (1985). Effects of microcomputer-administered diagnostic testing on immediate and continuing science achievement and attitudes. *Journal of Research in Science Teaching, 22*(9), 793-805.

Waugh, M. L., & Currier, D. (1986). Computer-based education: What we need to know. *Journal of Computers in Mathematics and Science Teaching, 5*(3), 13-15.

Willis, J. A. (1988). Learning outcome testing: A statewide approach to accountability. *Technological Horizons in Education, 16*(1), 67-93.