PlanAlyzer, an Interactive Computer-assisted Program to Teach Clinical Problem Solving in Diagnosing Anemia and Coronary Artery Disease

HAROLD C. LYON, Jr., Ed.D., JAMES C. HEALY, M.D., JAMES R. BELL, M.D., JOSEPH F. O'DONNELL, M.D., EDWARD K. SHULTZ, M.D., MARGARET MOORE-WEST, Ph.D., ROBERT S. WIGTON, M.D., FRANK HIRAI, and J. ROBERT BECK, M.D.

Just as many of the diagnostic techniques and therapeutic modalities in use in medicine 50 years ago are no longer appropriate, neither are many of the methods that are used to teach medicine. The medical student of today faces an information overload. The amount of knowledge that must be assimilated has expanded geometrically, to the point that a weekday edition of the *New York Times* contains more information than the average seventeenth-century professional was exposed to in his or her lifetime.[1] Although problem solving and decision analysis are assuming greater importance because of new technology, these skills are often neglected in favor of rote memorization of proliferating facts.[2]

Events in medical education during the past decade have created an impetus and demand for interactive computer-based, self-paced, case-based programs for medical students. Early in the decade, the "GPEP Report"[3] indicated this need in four of its recommendations and strongly encouraged the use of interactive, computer-based programs that teach the process of medical problem solving.

There is a pressing need for better preparation in hypothesis generation, clinical reasoning, the use of emerging interactive computer technology, and self-paced problem-based learning methods to help the student learn medical decision making under conditions of uncertainty, which are always present. Traditionally, "putting it all together" during the clinical experience is often the only instruction offered in the sequential process of decision making necessary for diagnosis and management of patients. This process is most often taught indirectly through the Socratic method (a euphemism for the process that occurs during ward rounds), or by simply patching together the algorithmic trees observed and experienced on the wards. This critical aspect of medical education is recognized in the recent attempts to introduce problem-based material earlier in the curriculum. Too often the teacher is unable to make clear and explicit to the medical student the path his or her mind has taken in diagnostic decision making.[4] In medical education, there is an ever-growing discrepancy between the continuously increasing multitude of facts and the low level of understanding of how to use them.

Even after two decades of research on medical problem solving, we still do not know precisely how the mind's iterative process leads to a diagnosis.[5] Barrows and Tamblyn[6] corroborate this view in their research, which suggests that accurate diagnostic labels often literally "pop" into experienced clinicians' minds within moments of their initial encounters with patients. The research does not even support a close and predictable connection between the accuracy of a physician's diagnosis and the quantity of the data collection. The overall quantity of findings recalled is not related to expertise,[7] though recall of significant findings may be.[8] Among those physicians who reach a correct diagnosis, there is a widespread tendency to overemphasize positive test findings while misinterpreting or ignoring negative ones and to deny findings that conflict with a pet hypothesis while obtaining redundant information to support the hypothesis.[8]

Hammond has suggested that the process of diagnosis always has "irreducible uncertainty."[9] Because of the irreducible uncertainty, diagnosis requires the exercise of judgment as well as knowledge.[10] Diagnostic skill seems to depend on three primary factors that help the individual clinician deal in individual ways with uncertainty: *experience*, which gradually is distilled into more and more abstract cognitive constructs; skill in the more content-oriented process called *pattern recognition*,[11] and skill in the more process-oriented process known as the *hypothetico-deductive process*, which facilitates clinical reasoning. Medical educators are still debating which of these is more important. We accept that skill in these three processes allows diagnoses to be made by experienced clinicians with sparse or incomplete information, whereas the novice requires all of the intermediate data that the experienced clinician assumes. We find that it is not easy for the clinician to convey these abstractions to students, and some evidence exists that physicians are fairly unaware of the process except in a superficial way. That thesis suggests that medical students of different experience levels are better taught by different methods. For example, neophytes are likely to be served best by what Friedman and associates call the "pedagogic" simulation format, which offers considerable support in accessing information and ample feedback on diagnostic options.[12]

We agree with some who maintain that computer technology has the power and capacity to facilitate independent learning and to teach problem solving in a manner unmatched by any medium other than real experience with patients.[3,13] Through carefully controlled studies, we must find cost-effective methods to achieve this. The PlanAlyzer project is an attempt to do this, and this paper presents some encouraging results from six years of development and research, including two years of carefully controlled trials. We discuss our experience with the PlanAlyzer project, future plans, and new research opportunities.

Background

PlanAlyzer was designed as a pedagological tool, taking into account several characteristics common to the clinician problem solver: limited capacity for short-term memory; use of heuristic strategies to examine promising avenues; a tendency to search for information sequentially; and the importance of the problem solver's conceptualization of the problem at hand.[14]

Heuristic strategies are methods of organizing knowledge and experience and are influenced by the idiosyncrasies of individual clinicians. Kassier and colleagues also note that "a large amount of knowledge, highly organized for the task at hand, is an essential ingredient in expert performance. Without such organized knowledge a person is forced by the limits of his cognitive abilities to pursue a rather plodding and often inefficient search for a solution to the problem before him."[14,15] The clinical problem solver suffers cognitive strain because of limited short-term memory and because information needed to solve the problem is not always accessible (e.g., forgotten information) or clear (e.g., an X-ray difficult to interpret). To cope with this cognitive strain, the physician organizes his or her knowledge base in special ways using heuristic strategies of elimination to limit the number of hypotheses being considered and thereby think more efficiently. Different settings, patients, and specialty types may combine to generate different types of heuristic strategies.[16] For example, the problem solving of hospital-based specialists is different from that of the primary care family physician. The goal of diagnostic reasoning in the specialties is to determine etiology and to assign an illness to a place in the taxonomy of disease. The goal of reasoning in family practice often is different. "The family doctor sees many patients with self-limiting disorders. In these a diagnosis is not required: it is sufficient to know that serious illness has been excluded."[17]

It is known that this process of medical problem solving involves a rapid gathering, organizing, and evaluation of clues in an intuitive process. From our research with PlanAlyzer, we believe that the processes of highlighting, linking, and integrating those clues to the diagnosis early in the presentation of a case are pedagogically important.

PlanAlyzer was conceived[18-20] to model the diagnostic processes of the expert clinician in order to serve as a point of departure for process-oriented instruction in medicine.[21,22] For this reason, PlanAlyzer presents sequential, highly organized case-based information to provide the student with simulated experience and with practice in the hypothetico-deductive process and pattern recognition. These goals are fostered by having the student practice diagnostic analyses that are patterned after expert clinicians' diagnostic strategies. These are sophisticated strategies, for the experienced clinician sees thousands of patients, and his or her knowledge base becomes "tuned" to solve clinical problems.[15] In this way, the design of PlanAlyzer focuses on the uniqueness of the clinical problems modeled and the domain experts' approaches to them. This approach has been echoed in related problem-solving research.[23,24]

History and Description

The project was conceived in 1985 by one of the authors (JR Beck) in the Dartmouth Medical School Program in Medical Information Science. It was funded by the National Library of Medicine (NLM) in 1986. The six-year PlanAlyzer project consisted of two years of initial design and development (1986–1987); a pre-trial formative evaluation year (1988), during which the entire class underwent a dress-rehearsal trial that resulted in significant developmental changes in the program; and two full trial years (1989–1990), during which all students in both the second-year classes were stratified by sex and MCAT scores and randomized into experimental groups (those using the computer-based program) and control groups (those using a traditional text version of the same cases) for the anemia course. Later each year there was a crossover, with the controls becoming the experimental subjects for a similar intervention with the cardiology course. At the end of each course, the students took posttests and completed detailed questionnaires. The final year (1991), during which all students used both the computer and text versions of the cases, was used for analysis of the research data.

The scientific goals of the project, "Computer-based Exercises in Pathophysiologic Diagnosis" were

1. To develop two computer-based laboratory exercises for basic content areas in medical education, oriented toward the processes of clinical reasoning in diagnosis, using techniques of medical decision science and software engineering.
2. To use the computerized teaching program to test the hypothesis that students with access to process-oriented educational tools can integrate their knowledge more effectively and/or efficiently than they can when they have access only to non-process-oriented traditional text-based educational materials.

System. The Apple Macintosh was selected for the user system, as it is widely available at Dartmouth, fairly inexpensive, and has a easy and intuitive user interface.

Design. For two years of initial design and development, the educator (HCL) worked closely with the two expert clinicians in cardiology (JR Bell) and anemia (JFO), who had been teaching these courses to second-year students as part of the Scientific Basis of Medicine (SBM) course, to determine and articulate what specific skills, concepts, and kinds of pattern recognition were necessary to diagnose chest pain and anemias successfully. This tedious but enlightening process revealed gaps in the past course content and resulted in curriculum changes even before development of the computer-based program. Of interest is the fact that one of the content experts (JFO) was not computer literate, but, as academic dean, nevertheless, felt responsible and highly motivated to make use of the emerging technology, which he now relishes using in his teaching. Working with the principal investigator (JR Beck) and the programmer (FH), the clinicians and the educator produced "gold standard" (i.e., derived from experts) work-up algorithms for each possible clinical situation for both coronary artery disease (CAD) and anemias. These were based on history, physical examination, and various test results. The team worked to obtain the specific knowledge (patterns and clues) and skills (hypothetico-deductive process) needed to identify the correct differential diagnosis for each variety of disease presented in the universe of cases. This content was reformulated to be programmed into sequential learning activities for students. Each of the cases has a "gold standard" work-up scheme that was encoded into the computer to give feedback to the student on tests ordered, sequence and appropriateness of testing, and performance characteristics of tests as the student gradually arrives at a diagnosis for each case. Each knowledge area required for differential diagnosis and management of the clinical problem was examined further to determine specific concepts, patterns, clues, vocabulary, and procedures the student would need to find in the patient's history, physical examination, and laboratory tests to make appropriate medical decisions to solve the patient's problem.[25,26]

Objectives. There are four instructional objectives common to each instructional area. At the end of the instruction the student should be able to (1) use the computer to complete the computerized cases; (2) extract the relevant information (using pattern recognition) from the patient's history and physical data to discern which laboratory tests are appropriate and cost-effective; (3) provide a supporting rationale, including an initial diagnosis, for laboratory test requests and give a plan for follow-up after results are obtained; and (4) interpret data to identify appropriate diagnosis and management strategy, based on the critique provided by PlanAlyzer.

The team developed 15 cases for each of the two content areas that illustrate the essential aspects of the universe of anemias and chest pain work-ups being taught. These cases include the required knowledge base and the decision process to appropriately address the patient's problem. These cases and their embedded knowledge bases and decision strategies serve as the basis for instructional content and assessment strategies. Pedagogically, principles of effective teaching were employed in the design of the program, including multisensory input from digitized sound and images, reviews, motivational steps, use of humor, "empathic" feedback, and critiques, where possible.

Students' use of the program. Students were encouraged to work on the cases in small groups in the library several days a week for seven weeks during each course until they completed all the cases, progressing from easy to more difficult ones. Typically, the students open the PlanAlyzer program on a Macintosh II computer with color monitor and are welcomed and introduced to the anemia (or CAD) case by the digitized voice (combined with graphics and text) of the content expert (JFO or JR Bell). The patient is briefly presented, and the students are immediately prompted to think about and type in what other history they might want. All responses are recorded by the computer in hidden files. A list of signs and symptoms (patterns or clues) is presented, and the students must click on each to learn whether the patient has or does not have each sign or symptom. At any time a student may "click on" a message button and send an e-mail message to team members. In addition, there are numerous teaching buttons that take the student to small multilevel tutorial modules such as one that offers the strategy of using a single laboratory test, the mean corpuscular volume (MCV), the results of which can eliminate about two-thirds of the possible anemias. Also, favorite quotes from the clinicians are available along with humorous cartoons in an attempt to give the cold computer technology some of the warmth and humanness found in the research to be present in successful teachers.[27] The student gathers data on the patient, creates a ranked or weighted hypothesis list (differential diagnosis), orders laboratory tests to test these hypotheses (including the making of an algorithmic tree for various test results), makes a diagnosis, and receives a critique of the entire process, generated by a comparison of the expert's algorithm and work-up with the student's. After a textual wrap-up discussion of the entire case, the student progresses to the next case.

Assessment. Prototype text cases and computerized cases were tested on sophomore medical students during the third year. The results were encouraging and students liked the interventions. There were several content areas crucial to successful solving of the cases—blood smear and electrocardiogram interpretations—that most students found too difficult. They also argued forcefully for replacement of some seminar and lecture time with this instructional approach, since they felt that they did not have adequate time to complete the case material.

Control intervention. The control intervention became text versions of the same cases presented to the students, prepared by the same clinicians who prepared the experimental material. Both control and experimental groups had an introductory lecture. The text-based casebook given to the control group contained the same cases and content in paper form as was provided to the experimental group in interactive computer form. A reference manual was provided with the control group's case material, whereas the same content in on-line help was provided to the students in the experimental group.

Wright-Giemsa. The pre-trial formative evaluation in 1988 revealed a weakness in our anemia PlanAlyzer paradigm. Second-year students have not received adequate instruction in the interpretation of peripheral blood smears, considered important to the evaluation of anemias. At that time, we used 35-mm slides of classic morphologic findings. What was lacking, however, was specific instruction in the systematic review and interpretation of the peripheral smear. To solve this problem, one of us (JCH) created a "smear generator" system that provides digitized color smears for each of PlanAlyzer's anemia cases, enabling the students in the experimental group to access the smears for each case in color directly from the computer. This new educational module, which we call "Wright-Giemsa,"[28] enables PlanAlyzer to simulate the process of teaching smear interpretation because it models the instructor as if he or she were at the second head of the microscope. A smear-authoring system was developed that facilitates the creation of various smears for additional cases. A cell-and-smear database is part of the program, with in-depth teaching material about cells available by "clicking" on them. This graphic instruction was popular with the students and enabled them to complete the cases.

Changes after formative evaluation. The following changes in the first pretrial during the third year (1988) were implemented for the full trials in 1989 and 1990:

1. The factual (multiple-choice) pre- and posttests were revised to include "higher-order" multiple-choice questions in an attempt to better measure concepts being taught by the PlanAlyzer cases.
2. The clinicians revised the cases to correct inadequacies and problems found by students (as recorded in the computer files or written in the casebooks). Cases were added where necessary to ensure that each concept or process that they wanted the PlanAlyzer cases to teach was covered.
3. The control group members were required to use their own individual case books (and reference manuals) *in the library only.* This ensured equal access to the teaching materials for the experimental and control groups.
4. The principal reasons that students did not use the cases extensively in the pretrial were that the cases were added on top of the original lectures and seminars (creating work overload) and that the students had little incentive to complete the cases.

Several decisions were reached to correct these problems. First, three two-hour anemia seminars (delivered by 12 faculty) and one two-hour lecture (a total of 74 faculty classroom hours) were eliminated for the 1989 intervention, giving the students that time to work on the PlanAlyzer and paper cases. Second, in cardiology, one two-hour seminar (delivered by ten faculty) and one two-hour lecture (a total of 22 faculty classroom hours) were eliminated. Third, students were told that the completion of the computer or the paper cases (depending on which they had been assigned) was required in order to complete the course.

The team was willing to make these compromises in the research design to ensure adequate time for students' use of the instructional materials and the overall success of the research project. For this reason, it was important for the content experts to create cases that covered the concepts taught in the seminars and lectures that were eliminated.

As reported previously,[26] in comparing the overall SBM scores for hematology and cardiology, we were encouraged to find that even after eliminating the traditional classroom hours of instruction mentioned above, the students scored slightly (though not significantly) higher on the examinations in 1989 with PlanAlyzer than they did in 1988 with the combination of PlanAlyzer and the traditional instruction.

5. The smear generator, as described earlier, and the EKG Tutor were introduced as interactive teaching modules to address second-year students' weaknesses in these areas. The equivalent textual content was added to the workbooks.

Evaluation

We carried out two types of evaluation: (1) formative evaluation, the dynamic process of making changes from pretrials during the development; and (2) summative evaluation, based on comparative differences in proficiency and efficiency from posttests, computer data, and questionnaires.

Efficiency and Economy Gains

The posttests at the ends of the 1988 pretrial and the 1989 and 1990 trials showed that the two programs had achieved the proposed instructional objectives, plus some significant efficiency and economic gains. Ninety-six (96) faculty hours of traditional direct student contact were eliminated by substituting PlanAlyzer in their place, while at the same time maintaining a high level of student achievement. Calculating an average of four hours preparation time for every classroom hour, these 96 hours, when calculated in dollars of faculty time saved, would be in the vicinity of $14,000 per year, which helps offset the initial start-up costs of approximately three of the Macintosh II workstations per year. In terms of students' proficiency, at the end of the instruction the 328 students in both the control and the experimental groups over both years were able to accomplish the instructional objectives.

One significant finding ($p < .001$ for both anemia and cardiology) was that the students in the experimental groups were able to complete their cases in 43% less time, with no loss in achievement, than the time needed by the students in the control group studying the same cases in a text-based version, as reported in 1991.[26] On the average, for both the anemia and the chest pain programs, students spent approximately 7.5 hours longer on the 30 cases on the workbooks than on the computer cases to achieve the same level of mastery. These data should be accepted with caution in that they are self-reported estimates from a questionnaire completed after the intervention. However, we would not expect such large student recall bias in remembering time spent. Although it was impossible to accurately corroborate the time spent on the workbooks (other than checking sign-out and sign-in times), it was possible to verify these estimates for the computer from the computer log files.

Proficiency

Differences between groups' pretest scores were tested using analysis of variance.[29] Differences between groups in all posttest scores were tested using analysis of covariance, adjusting for the pretest scores. We also performed a secondary analysis of differences in posttest scores, separately for both the pretest score and each student's Myers–Briggs personality type. All tests were performed at the .05 level of significance and were two-tailed. The experimental and control groups differed significantly only with respect to the amounts of time spent on the cases. The anemia experimental groups' total posttest scores for two years were marginally significantly different ($p = .049$); while no difference was found between the scores of the cardiology groups.

In both the 1989 and the 1990 trials, there was no significant difference in posttest scores (adjusted for the respective pretest scores) between the control groups and the experimental groups for either the cardiology or the anemia cases. However, the students were, in general, very favorably impressed with the cases as effective teaching tools. The final comparative trial of the first version of PlanAlyzer was completed in December 1991; the results confirmed the efficiency findings of the previous trials.

Beginning in 1992, PlanAlyzer became a standard part of the Dartmouth Medical School curriculum, with all students receiving both the computer-based PlanAlyzer program and the text-based workbook previously used by the control groups.

Cooperative Work in Other Institutions

PlanAlyzer programs have been distributed to approximately 20 other medical schools, including several in Europe and Canada, for trials in different settings. In the United States, a consortium of leading medical schools in the field of medical information science, called the North East Medical School Consortium (NEMSC),[30] has been formed to share medical education software and information about ongoing work, ideas, and approaches, to avoid having schools continue the costly practice of "reinventing the wheel" and developing programs in isolation from one another. Giving the initial development costs in money, personnel, and time to develop and evaluate effective interactive programs, such cooperative practices can lead to substantial savings and opportunities for medical education and curriculum reform. As the emerging technology moves toward more compatibility (e.g., the recent Apple–IBM cooperative agreements), such sharing will become even more feasible. Several projects are under way in collaborative relationships among the University of Munich, Hannover Medical School, Charles University, Prague, Charité Medical University, Berlin, Potsdam University, and Dartmouth Medical School.

In a six-month trial conducted with medical students at the University of Munich,[31] we found that German students—even those who on preintervention questionnaires indicated that their knowledge of English was not good—were able to perform the PlanAlyzer anemia cases satisfactorily. The fact that the English language appears to post no significant barrier for German medical students was an encouraging finding, as one of the authors (HCL) is sharing a variety of other interactive multisensory Teachware for trials in Germany and other European countries where the need is also documented.[32] Plans are under way with the Universities of Potsdam and Hannover to develop PlanAlyzer case-authoring tools that will link with artificial intelligence authoring tools designed for pattern recognition; this link could enable the program to extract from experts authoring cases their clinical reasoning strategies and patterns, thereby making PlanAlyzer "smarter" with each case entered or authored.

Importance of the Findings

Both groups of students (those with PlanAlyzer and those with the more traditional casebooks) mastered the basic facts and concepts very well in both content areas. After analyzing two full years of trials, we conclude that the originally proposed PlanAlyzer double cross-over trial paradigm had built into it "insurance" for pedagogical success in that both the experimental and control groups used two proven educational methodologies: case-based teaching and self-paced learning. (It is interesting to speculate what might have been the result of a hypothetical third group of students who would have been given only a list of course objectives. However, it was difficult enough assuaging the half of the class in the control group by telling them that they too would have their turn on the computer when the cross-over occurred).

The evaluation of the PlanAlyzer Programs is unique in two ways: first, it is one of few studies in a medical curriculum involving entire medical school classes randomized into control and experimental groups. Second, our study successfully controlled for the factors (listed below) that most often confound research on computer-assisted instruction (CAI), as reported by Clark in his comprehensive meta-analysis of the CAI research literature.[33]

1. The same instructional *methods* (self-paced, case-based) were employed in both PlanAlyzer experimental and control groups.
2. The same instructional *content* was employed in both the control and the experimental groups.
3. The same *level* of curriculum material was used in both groups.
4. Identical *effort* was expended in designing the groups' courseware.
5. The *same* teacher developed both the experimental group's material and the control group's material for each content area.
6. "Novelty effects" for short-term interventions (e.g., less than four weeks) is another factor offered to explain learning gains attributed to computer instruction.[34] In PlanAlyzer these novelty effects were not likely to be present since instruction occurred over a longer period of time (i.e., nine weeks)
7. Our computer group did not have more instructional support than did the controls.

Clark maintains that the motivation or effort students invest in CAI depends, in large measure, on their beliefs about how difficult it is to learn from computers. Belief that computers are either very easy or very difficult may result in less persistence at learning tasks.[35] However, results from our questionnaires so far do not support his hypothesis, as there was no significant difference in the performances of those students with high and low expectations about the computer. (The 27 students with the highest expectations averaged 82.0% on the posttest compared with 83.5 for the ten with the lowest expectations.)

Both the PlanAlyzer cases and the text casebooks were in the domain of self-paced and case-based learning. In both the computer and the casebooks, the "expert" clinicians' step-by-step strategies were modeled into teaching vehicles (with the same content) that led the students through progressively more difficult cases. The only difference in the methods for the two groups was that one was more interactive than the other, which resulted in significant efficiency gains. This leads one to speculate: if the students on the computer were to use the 7.5 hours they saved to study or interact with additional materials or cases, would that additional learning result in better posttest scores than the scores of the controls, who spent so much more time on the cases? This line of questioning supports creating a larger base of cases, perhaps more than the slower group would be able to complete in the seven-week course. The efficiency advantages of the computer might then be translated into proficiency gains, since the subjects in the experimental group could use their saved time to complete more cases and consequently gain more experience in pattern recognition and clinical reasoning skills. Research on gifted learners supports the hypothesis that, given additional time, proficiency gains would result.[36] Although we found no significant proficiency difference between the experimental and the control groups who used the cardiology program and only marginally significant differences between those who used the anemia program, we are encouraged by Friedman and colleagues, who suggest that outcomes worthy of study include areas where differences are not expected: "While there is a tendency to give second-class status to studies with 'no-difference' findings, studies of medical problem solving and education are very important in the comparison of traditional and innovative programs of medical education."[37]

Also, we are concerned that our case-based paper posttests do not adequately measure the integrative process of clinical reasoning that PlanAlyzer was designed to teach—the complex processes involved in the integration of factual and conceptual knowledge, and the content in pattern recognition.

Other data from the students' questionnaires are summarized below:

1. The students preferred the computer over the casebooks, but liked both formats much more than the traditional instruction that the intervention replaced. Many recommended the use of both in future instruction.
2. More students had expectations that they would probably learn more from the computer than from the casebooks. However, there was no significant difference in the performances of those with high and low expectations of computer learning, as Clark reports in other studies.
3. The students felt the computer probably prepared them better as future physicians for chest pain and anemia diagnosis, for their examinations, and for their clerkship experiences.
4. The students believed that the computer program would allow them to retain what they had learned longer than would the workbook program.

Conclusions

The process of research and development of PlanAlyzer has been fruitful, especially from the standpoint of developing a promising curriculum intervention for anemia and CAD diagnosis that is popular with both faculty and students and that demonstrates significant efficiency gains (time savings) for both students and faculty. The emerging research in cognitive psychology—such as our collaborative work with Dr. Heinz Mandl from the University of Munich, Germany—offers new promise for a better understanding of the process of medical problem solving. Although the PlanAlyzer research has not yet contributed significantly to that understanding, PlanAlyzer has helped us understand the efficiency of computer-based systems in medical students' education. Also, since the computer captures a record of each student's problem-solving strategies, we have gathered an untapped database for further analysis and future research. The positive attitude of our students toward the computer intervention is consistent with those of other studies concerning students' acceptance of computer-based programs.[38] While findings of a more significant increase in posttest performance would be positive for any learning vehicle, most carefully controlled studies fail to show such differences, and it should not be considered the only appropriate factor in judging the success or failure of an educational approach. Many of the advantages of computer-based learning in PlanAlyzer—e.g., realistic interactive practice in the hypothetical deductive process and pattern recognition, individualized self-paced, case-based instruction, self evaluation, opportunities for collaborative learning in small groups —represent positive returns, perhaps even more important to us pedagogically than more significant posttest achievement gains.

Our findings have led us to understand that we need more effective instruments to better test student problem-solving ability. We propose an additional posttest in the form of a video-based multistep modified essay question (MEQ) exam where the students in an auditorium setting will be presented with patients by video in a *start* video–stop–answer short MEQ essay question–*resume* video–stop–answer question format. We believe that this will present realistic cases via video, while correcting for the deficiencies reported in the literature in similar objective structured clinical examinations and MEQ-type examinations.[39] We postulate that such examinations may better establish whether the students using PlanAlyzer are learning "problem solving" more effectively and/or efficiently than are the students with the casebooks.[40]

Clark[35] concludes that the computer *itself* does not make any difference in students' performances. The computer or the text is a vehicle to deliver the content. We have found that, in the same way that a drug can be delivered faster (more efficiently) by a hypodermic needle than by a pill, a properly designed computer program can be a more efficient delivery vehicle than the traditional text. However, the computer itself and associated hypermedia appear to offer a variety of other advantages besides efficiency gains, including better ways to organize information for the teaching of pattern recognition[41] and problem solving, advantages to certain types of students, speed, and divergent thinking advantages for some students. The factors listed earlier that we so successfully controlled for are the variables that, when uncontrolled, often produce findings that computer-assisted instruction leads to greater proficiency gains.

We believe with Clark that an appropriate shift in the paradigm for research in educational technology is overdue. This needed shift is from the traditional behavioral-outcomes approach to a cognitive-science approach where the focus is more on individual learner differences than it is on group posttest scores. Learners are multidimensional: some are more visually oriented, while others are verbally oriented. We need to begin to develop more intelligent CAI that has the capacity to respond to such learner differences. But first we must develop a better cognitive understanding of what those differences are and how to model learning options for different types of learners. Dr. Mandl in Germany is pioneering in such studies and we are working with him to design studies, using PlanAlyzer, of individual learners' differences.[42]

Investigators are presently at an interesting place in the evolution of educational technology, where the technological advances, particularly with the advent of interactive multimedia, are considerably ahead of our understanding of how the technology contributes pedagogically to the learning process. New advances in the productive field of cognitive psychology will help us understand the processes we are now experimenting with from the learners' perspective as we expose learners to the intuitive process of exploring, navigating through learning "landscapes" containing multilayered hierarchies of data and knowledge. Those who are developing such medical learning environments are pioneers and discoverers in new learning microworlds. Hypermedia provides a unique opportunity to both present and discover knowledge in ways that transform the teaching–learning process into an adventure for the learner, the teacher, and the developer.

We found with PlanAlyzer that the process of developing teaching software can dramatically transform the content that is being taught as well as the teacher's approach to teaching it. The dynamic process of developing interactive materials fosters effective diagnosis and new understanding of what is and is not being taught—a process that leads to important opportunities to correct deficiencies and improve the teaching. For example, we had not expected or planned to develop as peripherals to PlanAlyzer tutorial modules such as the Wright-Giemsa Smear Tutor, the EKG Tutor, and Bayes Calculator, which the process and formative evaluation led us to do.

In the larger context, such problem-based, self-paced teachware can become powerful vehicles for curriculum reform, promoting what the GPEP study says is needed: self-paced, problem-based, interactive learning that can help diminish the traditional gap between the scientific basis of medicine and real clinical practice. However, the introduction into the curriculum of such interactive medical education programs is not a trivial matter. Sometimes the initiative for using these programs comes from a generation younger than most faculty—students who were brought up learning from "Sesame Street" and other electronic media. In fact, a phone call to a particular large eastern medical university asking for the department of medical informatics will lead the caller to a student who has taken it upon himself to import interactive teachware from all around the United States and tactfully share it with selected faculty members, many of whom are now using it in their classes.

At Dartmouth we had several prerequisites for a successful trial: senior faculty (the associate dean for academics was one of our content experts) who supported the project; several junior faculty with considerable motivation; an experienced educator to ensure that the project was sound pedagogically and that the investment in the technology would result in a program that taught what we expected it would; and several years of funding from the NLM, without which we would not have been able to do the development or the study.

The new cognitive paradigm we are moving toward in computer-based research "assumes that instructional powers do not reside solely in the media, for the way we perceive media influences what we learn from them." On the other hand, "learners are not the sole power brokers, for their perceptions are founded on the kinds of information and instructional methods delivered by different media."[35] So we must look carefully at outcomes of interventions with the emerging technology; but we must also begin to focus more on the characteristics of the individual learners and how they perceive and interact with the technology. It is essential, as we expose medical students to our experiments, that we gather the data necessary to enable and facilitate the evaluations so necessary for determining what we are teaching and how effectively it is taught. We profess a "scientific" tradition that is based upon knowledge. The root of that tradition is understanding what occurs so that we can explain it to others and ourselves. Yet, in spite of our professed rationalism, we tend to promote promising innovations and design poor research to study them, the results of which are open to interpretation in a number of ways, one of which usually shores up our faith in our own innovation. We continue to do this, failing to explain why we are getting the results we observe, until the innovation begins to fail us, or like Toad in *Wind in the Willows*, we are confronted with a more glamorous innovation. We then drop the old innovation, having learned nothing from it, and begin the whole process over again with the next. This has been the history of innovations in the educational technology from radio through intelligent tutoring systems. It is time to break the cycle and do the careful, tedious, and often time-consuming research needed to find why we are getting the results we are.

This research was supported by grants LM04487, LM00086, and LM07044 from the National Library of Medicine and by a grant from the National Fund for Medical Education.

Dr. Lyon is assistant professor, Department of Pathology, Department of Community and Family Medicine, and the Program in Medical Information Science (PMIS); Dr. Bell is associate professor, Department of Medicine; Dr. O'Donnell is professor, Department of Medicine, and associate dean for student affairs; Dr. Shultz is associate professor, Department of Pathology; Dr. Moore-West is research associate professor, Department of Community and Family Medicine; and Mr. Hirai is senior programmer with the PMIS; all at Dartmouth Medical School, Hanover, New Hampshire. Dr. Healy is assistant professor, Department of Pathology, Creighton University School of Medicine, Omaha, Nebraska; Dr. Wigton is professor of internal medicine and associate dean for academic affairs and graduate medical education, University of Nebraska Medical Center College of Medicine, Omaha, Nebraska; and Dr. Beck is professor, Department of Pathology, and vice president for information sciences, Baylor College of Medicine, Houston, Texas.

References

1. Technology and the American Economic Transition: Choices for the Future. Office of Technology Assessment, U.S. Congress, Washington, D.C.: Government Printing Office, May 1988, pp. 240–251.
2. Piemme, T. E. Computer-assisted Learning and Evaluation in Medicine. *JAMA* **260**(1988):367–372.
3. Muller, S. (Chairman). Physicians for the Twenty-First Century. Report of the Project Panel on the General Professional Education of the Physician and College Preparation for Medicine. *J. Med. Educ.* **59**, Part 2 (November 1984).
4. Gremy, F. The Future of Information Processing in Medicine and Public Health. *Comput. Progr. Biomed.* **11**(1980):71.
5. McGuire, C. H. Medical Problem-solving: A Critique of the Literatures. *Res. Med. Educ. Proc.* **23**(1984):3.
6. Barrows, H. S., and Tamblyn, R. M. *Problem Based Learning—An Approach to Medical Education.* New York: Springer, 1980.
7. Muzzin, L. J., Norman, G. R., Feightner, J. W., Tugwell, P., and Guyatt, G. Expertise on Recall of Clinical Protocols in Two Specialty Areas. *RIME Proceedings* **22** (1983):122–127.
8. Patel, V. L. Expert Diagnostic Reasoning: An Interactive Process of Case Comprehension. *PERN* **5**(1984):7–11.
9. Hammond, K. R. Computer Graphics as an Aid to Learning. *Science* **172**(1971): 903–908.
10. Wigton, R. S., Poses, R. M., Collins, M., and Cebul, R. D. Teaching Old Dogs New Tricks: Using Cognitive Feedback to Improve Physicians' Diagnostic Judgements on Simulated Cases. *Acad. Med.* **65**, Supplement (September 1990):S5–S6.
11. Papa, F., Shores, J. H., and Meyer, S. Effects of Pattern Matching, Pattern Recognition, and Experience in the Development of Diagnostic Expertise. *Acad. Med.* **65**, Supplement (September 1990):S21–S22.
12. Friedman, C. P., France, C. L., and Drossman, D. A. A Randomized Comparison of Alternative Formats for Clinical Simulations. In *Proceedings of the 13th Annual Symposium on Computer Applications in Medical Care.* Los Angeles, California: IEEE Computer Society, 1989, pp. 893–897.
13. Greenes, R. A., and Shortliffe, E. H. Medical Informatics: An Emerging Academic Discipline and Institutional Priority. *JAMA* **263**(1990):1114–1120.
14. Kassirer, J. P., and Gorry, G. A. Clinical Problem Solving: A Behavioral Analysis. *Ann. Intern. Med.* **89**(1978):245–255.
15. Miller, P. L. Critiquing Anesthetic Management: The ATTENDING Computer System. *Anesthesiology* **58**(1983):362.
16. Bauer, L. C. Human Problem Solving in Clinical Medicine. Unpublished Manuscript, Milton S. Hershey Medical Center, Hershey, PA, 1986.
17. McWhinney, I. R. Problem Solving in Medical Practice, *Proc. R. Coll. Med.* **65** (1972):924–938.
18. Beck, J. R. Computer-based Exercises in Diagnosis: Interim Evaluation of the PlanAlyzer Project. In *Proceedings of the 14th Annual Symposium on Computer Applications in Medical Care.* Los Angeles, California: IEEE Computer Society, 1990, pp. 483–487.
19. Beck, J. R., O'Donnell, J. F., Hirai, F., Simmons, J. J., Healy, J. C., and Lyon, H. C. Computer-based Exercises for Anemia Diagnosis (PlanAlyzer). *Meth. Inform. Med.* November 28, 1989, pp. 364–369.
20. Lyon, H. C., et al. Computer-based Exercises in Anemia and Chest Pain Diagnosis: An Interim Evaluation of the PlanAlyzer Project. In *Proceedings of the 14th Annual Symposium on Computer Applications in Medical Care.* Los Angeles, California: IEEE Computer Society, 1990, pp. 483–487.
21. Berner, E., Bligh, T. H., and Guerin, R. O. An Indication for a Process Dimension in Medical Problem-solving. *Med. Educ.* **11**(1977):324.
22. Patel, V., HoPingKong, H., and Mark, V. C. Role of Prior Knowledge in Comprehension of Medical Information by Medical Students and Physicians. *Res. Med. Educ. Proc.* **23**(1984):127.
23. Joseph, G. M., and Patel, V. L. Domain Knowledge and Hypothesis Generation in Diagnostic Reasoning. *Med. Decis. Making* **10**(1990):31–46.
24. Elstein, A. S., Shulman, L. S., and Sprafka, S. A. *Medical Problem Solving, An Analysis of Clinical Reasoning.* Cambridge, Massachusetts: Harvard University Press, 1978.
25. Lyon, H. C., et al. Findings from an Evaluation of PlanAlyzer's Double Cross-Over Trials of Computer-Based, Self-Paced Case-Based Programs in Anemia and Chest Pain Diagnosis. In *Proceedings of the 15th Annual Symposium on Computer Applications in Medical Care.* Los Angeles, California: IEEE Computer Society, 1991, pp. 88–93.
26. Lyon, H. C., et al. Significant Efficiency Findings while Controlling for the Frequent Confounders of CAI Research in the PlanAlyzer Project's Computer-Based, Self-Paced, Case-Based Programs in Anemia and Chest Pain Diagnosis. *J. Med. Systems* **15**(1991):117–132.
27. Carkhuff, R. *The Productive Teacher.* Amherst, Massachusetts: Human Resources Development Press, 1984.
28. Healy, J. C., Rainer, R. O., and Beck, J. R. Wright–Giemsa: Automated Generation of Peripheral Blood Smears for Hematology Education. In *Proceedings of the 13th Annual Symposium on Computer Applications in Medical Care.* Los Angeles, California: IEEE Computer Society, 1989, pp. 907–910.
29. Neter, J., Waserman, W., and Kutner, M. H. *Applied Linear Statistical Models, 2nd Edition.* Homewood, Illinois: R. D. Irwin, 1985.
30. Alonzo, D., et al. NEMSC Cooperative Development of Medical Courseware (Abstract). In *Proceedings of the Association of American Medical Colleges,* October 1990. [The NEMSC includes Cornell, Dartmouth, Georgetown, Harvard, Johns Hopkins, Mt. Sinai, New York University, SUNY, the University of Pennsylvania, the University of Rochester, Yale, Stanford, and Jefferson Medical University.]
31. Mandl, H., Graesel, C., Prenzel, M., Bruckmoser, J., Lyon, H. C., and Eitel, F. Clinical Reasoning in the Context of a Computer-based Learning Environment. Research Report No. 4. Munich, Germany: Ludwig-Maximilians-University, Institute for Empirical Pedagogy and Psychology, December 1991.
32. Das Arztbild der Zukunft Analysen künftiger Anforderungen an den Artz Konsequenzen für die Ausbildung und Wege zu ihrer Reform, Abschlussbericht des Murrhardter Kreises, Robert Bosch Stiftung, 1989.
33. Clark, R. E. Confounding in Educational Computing Research. *J. Educ. Comput.*

Res. 1(1985):137-148.
34. Clark, R. E. Research on Student Thought Processes in Computer-based Instruction. *J. Instr. Devel.* 7(1984):2.
35. Clark, R. E. Research on Instructional Medial 1978-1988. In *Educational Media Yearbook, 1988*, D. Ely, ed., pp. 20-35. Denver, Colorado: Libraries Unlimited, 1988.
36. Lyon, H. C. Our Most Neglected Natural Resource—A Special Feature on the Gifted and Talented. In *Today's Education*, February-March 1981, pp. 15-20. Washington, D.C.: NEA Publishing, 1981.
37. Friedman, C. P., et al. Charting the Winds of Change: Evaluating Innovative Medical Curricula. *Acad. Med.* 65(1990):8-14.
38. Walsh, R. J., and Bohn, R. C. Computer-assisted Instructions: A Role in Teaching Human Gross Anatomy. *Med. Educ.* 24(1990):499-506.
39. Rabinowitz, H., and Hojat, M. A Comparison of the Modified Essay Question and Multiple Choice Question Formats: Their Relationship to Clinical Performance. *Fam. Med.* 21(1989):364-367.
40. de Graaff, E. Test of Medical Problem-solving Scored by Nurses and Doctors. *Med. Educ.* 23(1989):381-386.
41. Norman, G. R. Simulated Patients. In *Assessing Clinical Competence*, Neufeld, V. R., and Norman, G. R., eds. New York: Springer, 1987.
42. Mandl, H., and Fischer, P. M. *Hypermedia and Learning*. Cognitive Science Series. Berlin, Germany: Springer-Verlag, 1991.

JOHN C. ROSE, M.D.
Associate Editor

Book Reviews

The Medical School's Mission and the Population's Health

Edited by Kerr L. White and Julia E. Connelly. New York: Springer-Verlag, 1992, 281 pages, $59.00.

This book is the proceedings of a conference held in December 1990. Although it suffers from differences in the styles of the various writers, it should be read by everyone concerned with the role of medical schools in our current and future society. The principal theme is that medical schools must adopt a mission that includes meeting the "individual and collective needs of society." Clearly, that will be accomplished primarily through the schools' educational activities, by preparing medical students and residents to provide appropriate care for patients but also for communities.

The first chapter, by White and Connelly, sets the theme for the rest of the chapters, and it is best summarized by the following quotation: "Society's expectations, political views of the role of government, vested interests, administrative strategies, and professional intransigence all contribute [to the problems of health care] . . . . For this state of affairs, the academic branch of the profession must assume much of the responsibility and now has the opportunity to provide renewed leadership." They also emphasize that there are some contemporary developments favoring change, including the "information revolution," the need for better communication, molecular biology, concern with the environment, and the managerial era. They end with the question, "what then is the mission of the medical school?"

Two specific examples of changes in mission are described in some detail. The first is that of a "traditional" school—Albert Einstein—and the second that of a "community and population oriented" school—Newcastle, Australia. The latter is a new school, having been founded in 1974, with the goal of developing generalists. The curriculum as adopted in 1990 is built around five "Domains of Learning" of which Domain III, Identification, Prevention, and Management of Disease, is the main organizing feature. It will be interesting to see future reports of the success of this approach. The Einstein program grew from a Department of Ambulatory Medicine established at the North Central Bronx Hospital (NCBH). That department has established research, patient care, and education programs related directly to the community served by the hospital. Einstein has established a required two-month clerkship in ambulatory medicine in the fourth year as well as required ambulatory experiences in the third-year clerkships. The students assigned to the NCBH are immersed in the programs of the Department of Ambulatory Medicine.

One chapter deals with the establishment and implementation of entities known as "Health Intelligence Units" designed to monitor the health status of a defined community and to assess the need for specific educational programs. Specific examples, such as the Health Priorities Analysis Unit at McMaster University and the Community Health Research Unit at the University of Ottawa, are described to illustrate the relationship between educational programs and population-based data collection and analysis. Special emphasis is given to the possibility of critical linkages among the university, the community served, and the health care system generally.

Other chapters discuss the essential institutional competencies for population-based education and the essential population-based competencies for medical students. Such competencies include preventive medicine, knowledge of the organization and availability of health-related services, and descriptive epidemiology.

Medical schools exist to serve society, and, with the growth in knowledge and societal expectations, their missions must expand to include a broader range of functions. These authors make a very strong case for community-oriented education and research programs integrated with those for health care. They recognize that the primary mission of medical schools is education, but they emphasize that such education must include competencies in the skills necessary for care of a community. To repeat, anyone interested in the roles of medical schools should read this book.

Edward N. Brandt, Jr., M.D., Ph.D.

Dr. Brandt is professor, Department of Health Administration and Policy, College of Public Health, and is co-director, Center for Health Policy Research and Development, Colleges of Public Health and Medicine, University of Oklahoma Health Sciences Center.