THE EFFECTS OF A TUTORIAL AND A PROBLEM-SOLVING APPROACH ON THE PERFORMANCE OF MEDICAL STUDENTS: A COMPARISON OF TWO COMPUTER BASED INSTRUCTIONAL STRATEGIES

by

Evie Tsouna-Hadjis

A THESIS PRESENTED
TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
McGILL UNIVERSITY IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY AND SOCIOLOGY
MARCH, 1980

Résumé

Cette recherche a porté sur l'efficacité de deux méthodes d'enseignement automatisé par l'ordinateur dans le cadre d'une école de médecine. Des éetudiants de deuxième et quatrième années ont été exposés soit à un programme d'enseignement individuel, soit à un programme des simulation, ces deux programmes ayant été conçus par la même personne (un pharmacologiste clinique) pour enseigner le même sujet: soit la façon de s'occuper d'un malade souffrant d'un empoisonnement à l'aspirine. Le rendement a été évalué en fonction des comparaisons pré-test et post-test. Chaque test comprenait deux composantes permettant d'évaluer la connaissance pertinente des sciences de base et des méthodes cliniques. L'auteur a découvert que le programme d'enseignement individuel produisait un rendement bien meilleur aussi bien sur le plan théorique que sur le plan clinique. Il n'a pas trouvé de différence entre le rendement des étudiants de deuxième année et celui des étudiants de quatrième année, ni dans le pré-test ni dans le post-test. Par ailleurs, une comparaison entre les étudiants de quatrième année à rendement supérieur et à rendement inférieur a révélé que les étudiants à rendement inférieur tiraient bien plus de profit du programme d'enseignement individuel.

ACKNOWLEDGEMENTS

The author wishes to express her deep appreciation to all the people who in some manner assisted in the carrying out of this project.

I am particularly grateful to my advisor, Dr. Guy J. Groen, for his most essential contribution to the completion of this project, and for his kind willingness to offer his time, knowledge and interest. Also, special thanks are due to Dr. J. H. Kreeft for the many hours of hard work on the experimental material and for his valuable cooperation.

I am also most grateful to Dr. H.M. Scott, former Director of the Centre for Medical Education, for his approval, encouragement and financial support.

A special acknowledgement is given to Avril Bray for her continuous encouragement and assistance and for administrating the successful typing of the manuscript by Kathryn Boyd and Elly Jenkins.

Finally, a special note of thanks to all the medical students who enthusiastically participated and contributed to the success of the experiment.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

APPENDICES

Appendix | Page No.

CHAPTER I

REVIEW OF THE LITERATURE

This research is concerned with the effectiveness of two computer-based instructional strategies in a medical school setting: a tutorial strategy and a problem solving strategy. This issue is closely connected with a classical area of research in educational psychology: the relative effectiveness of expository and discovery methods of teaching. The present chapter will begin with a review of relevant aspects of this literature. This will be followed by a review of two other areas: medical problem solving and computer based instruction in medical education.

Two General Instructional Methods

Studies on Their Effects on Learning

In the field of educational research, the role of the two major teaching methods, loosely referred to as 'expository' and 'discovery' teaching, has been the object of many studies and, not infrequently, of controversies. In effect, one of the arguments accompanying the former method is that teaching material must be presented deductively in small, explicit parts--an approach that implies guidance and tutoring in the teaching act. On the other hand, a major argument in the latter method is that a loose teaching environment is most effective, as it may enhance spontaneous learning and discovery on the part of the learner. Thus, learning by discovery and problem solving is emphasized. Quite often, the above two teaching methods are used together in a combined form. This third method is usually called 'guided discovery'.

The field of research on the relative effectiveness of the above instructional approaches is rich with empirical studies the findings of which have shed some light on such variables as initial learning, retention, and transfer of subject matter. Due to a certain complexity of the area with respect to tasks and approaches which are not always well defined, these findings are sometimes contradictory, especially those of the earlier studies. For example, studies carried out by Hendrix (1947) and Gagne and Brown (1961) gave support to a significant superiority of the 'discovery' method over the 'expository' method with respect to transfer of learning, and a study by Ray (1961) revealed the same superiority with respect to retention. However, other studies by Craig (1956) and Kittel (1957) found the same kind of superiority for the 'expository' method. More recent studies have tended to support the former viewpoint about the 'discovery' method, yet reserve other properties for the 'expository' one. Nowadays, the differences between the two major methods are believed to be qualitative rather than quantitative, a fact that leads to an increasing interest in the combined method of 'guided discovery'. Wittrock (1963) first pointed out that subjects who learn by means of different procedures perform different responses which imply different outcomes of learning. In his experiment he tested the outcome of the three teaching approaches in terms of initial learning, retention and transfer. He found that maximum amount of direction produced the greatest initial learning, minimum direction was least effective on the three variables and finally, intermediate amount of direction produced the greatest retention and transfer. Guthrie (1967) came up with very similar results, namely, that the discovery method appears to facilitate transfer but not retention while expository instruction facilitates retention but impedes transfer. As Guthrie's data reveals, the more independent a teaching method is the more the transfer that can be achieved. Moreover, the amount of training required for the two approaches seems to be different, with the expository instruction requiring the less time. A study by Worthen (1968) showed that an expository method led to slightly better performance on an immediate achievement test covering the concepts actually taught. However, the students taught by the discovery method remembered the concepts slightly better when the test was repeated at a later date. In another interesting experiment by Maier and Burke (1966) problem solving resulted in general transfer of learning while guidance with a construction similar to that needed for the solution of the problem, aided performance of the initially unsuccessful subjects. This study also illustrated the value of both approaches with respect to different subjects.

It is apparent in the above studies that expository teaching tends to be superior for achieving immediate learning, simply because, as Shulman (1974) suggests, the subject-matter concepts are emphasized more. On the other hand, discovery seems to be superior for the development of broad inquiry and competency but not for the specific concepts. Actually, these concepts seem not to be learned well because they would interfere with remote transfer (Shulman, 1974).

Unfortunately, the findings of the above studies cannot serve yet as the elements of prescriptions for teaching strategies. They are much too general, often not too reliable and it is very unlikely that they hold true for all the task categories that education deals with. In effect, these findings can only serve as stepping stones toward more systematic research with respect to the particular task of different educational fields. Evidence relevant to this comes from a study by Goldman and Hudson (1973) in which it was found that students use different approaches for the completion of tasks from different educational fields. At this stage then, educational research seems to be in need of evidence from studies involving tasks from specific educational fields. In this way, familiarity with task specific variables may, in turn, make possible an approach of generalization among fields with more chances of success.

Medical Education and Clinical Problem Solving

The field of medical education offers to educational research a variety of tasks which involve not only pure factual knowledge but also skill in problem solving. In particular, clinical problem solving (i.e. the diagnosis and management of the patient) requires factual knowledge and clinical skills to be used by the student in a coherent and effective way.

Several attempts have been made to define the overall process of medical problem solving (e.g. Hoffman, 1960;

Shulman, Loupe and Piper, 1968; Elstein, Kagan, Shulman, Jason and Loupe, 1972; Barnoon and Wolfe, 1972). A complete review of this extensive work will not be attempted here. The discussion will be restricted to a brief review of those aspects most closely connected to modern cognitive psychology (Elstein, Shulman and Sprafka, 1978; Barrows, Feightner, Neufeld and Norman, 1978). From this point of view, the process of problem solving begins wih an adequate understanding of the problem. This involves an adequate state of knowledge (Maier, 1970) which can be used for the interpretation of the information given in the problem. This implies, of course, a knowledge of the subject matter involved in the problem. It also implies the possession of a set of procedures or skills for solving the problem. To quote Simon (1977): "Knowledge without suitable procedure for its use is dumb, and procedure without suitable knowledge is blind". Simon also suggests that understanding the problem gives rise to the next stage in the process, which is the formation--or retrieval--of a plan. The plan or hypothesis in which the goals and means towards the solution are defined (Miller, Galanter and Pribram, 1960) is followed by subsequent plans and subplans of proceeding. A general heuristic for accomplishing this is to evaluate the difference between the present situation and the goal and to reduce it one step at a time. Newell and Simon (1972) call this procedure means-end analysis.

The main concern of this thesis is with so-called patient management problems, in which the physician is presented with a patient with a given set of symptoms and the task is to provide an accurate diagnosis followed by a satisfactory therapeutic procedure. The physician, after interpreting the information given in the problem, forms his plans on how to proceed. The normal strategy is to obtain more data from the patient (by physical examination and/or laboratory tests), make an initial diagnosis and then manage the patient. Once the patient is out of immediate danger, more data can be collected (Johnson, Mollez and Bass, 1975). Apart from this overall approach, a number of other decision points in the development and execution of a plan are required. For example, he may have to decide which of the possible sources of information about the patient to use, when laboratory tests should be asked for, what particular treatment would be more effective and, in general, he must manipulate in an effective way the many interrelated variables. In other words, he must be able to confirm or disconfirm his hypotheses on time.

Results by Elstein _et al_ (1972,1978) suggest that, in a clinical problem solving situation, the generation of diagnostic hypotheses from the physician's background knowledge tends to occur very early in the problem solving process. Moreover, the size of the pool of available hypotheses tends to be quite small (normally, six hypotheses or less). Barrows et al (1978) have replicated and extended these findings in a series of studies of the process of medical diagnosis. Their main additional findings are: (a) the hypothesis pool remains stable after the first quarter of the encounter; (b) the clinician appears to be able to simultaneously process about three hypotheses on the average;

(c) information-gathering is strongly influenced by a search for data in support of the hypotheses. Curiously enough, they find that the process of clinical problem solving remains relatively constant across educational levels. However, the contents of the hypotheses do change as a function of educational level.

In their recent book on medical problem solving, Elstein, Shulman and Sprafka (1978) explain these and related results on the basis of the Newell-Simon theory of problem solving. They do this in the context of a general model for medical inquiry that encompassess four major processes--cue acquisition, hypothesis generation, cue interpretation and hypothesis evaluation or judgement. Hypothesis generation involves retrieving a limited number of hypotheses from long term memory and setting them up as a problem space. Cue acquisition depends heavily on routinized knowledge of history taking and routine physical examination.

The initial processing that occurs in a clinical situation is compared by these authors to the way a chess master reasons about chess. De Groot (1965) found that chess grand masters were not distinguished from weaker players in planning ahead or thinking more deeply. The only differences he could identify were in memory and perception. The data considered by Elstein, Shulman and Sprafka leads to the same conclusion, which also holds true in other areas of realistic problem solving such as logic and physics. The differences between experts and weaker problem solvers are more to be found in the repertory of their experiences, organized in long term memory, than in differences in the planning and problem solving heuristics employed.

To account for the phenomenon of early hypothesis formation, Elstein, Shulman and Sprafka make use of two fundamental propositions from the Newell-Simon theory: (a) that the task environment (i.e. the problem) is represented internally as a problem space, and (b) that the structure of the problem space determines the process to be used in the search for a solution to the problem. In clinical medicine, the potential size of the problem space is likely to be very large. It is necessary to find some way of limiting the size of the space to be searched. The early generation of tentative diagnostic hypotheses is a major strategy used by clinicians to define the regions of the potential problem space most likely to yield the solution. The method used to narrow these hypotheses or select a therapy is a form of means-end analysis in which specific clinical findings or clusters of findings serve to reduce the distance between the point where the problem solver is and the final goal.

Computer Based Instruction in Medical Education

As a number of authors have pointed out, Medical Education has been faced with an increased demand for more physicians together with a need for efficiency in coping with the growing body of medical knowledge (Stolurow, 1970). These reasons have made the need for individualized instruction more compelling, especially in the use of computers since this medium can give more information to students in less time, on an individual basis. The usual medical curriculum involves the teaching of basic medical material (i.e. physiology, anatomy, etc.) in lectures, during the first two years of medical study and the teaching of clinical competency within the third and fourth year by means of practice in the hospital. Individualized instruction is considered capable to aid the teaching of both basic knowledge and clinical skill.

Tutorial individualized instruction, frequently computer-based, has been used for many years in the medical curricula and by tradition applies mostly to the teaching of basic factual knowledge. On the other hand, the problem solving approach is a fairly recent application which is used mainly as a means towards clinical competency. The devices that are used nowadays as supplements to realistic clinical practice are simulations of clinical problems, not necessarily computer-based. These are intended to be very realistic presentations of the problems usually encountered by the physician in the hospital. One purpose for this is motivation and direct transfer to the reality of the hospital. Another is to allow the student to learn about procedures in which a real-life mistake would have disastrous consequences.

<u>Major Applications of Tutorial Computer-Based Instruction</u>

Tutorial CAI has been used within the last decade in various forms by a number of medical schools. The most extensive application of educational technology is at the University of Illinois, College of Medicine. In this school, 400 self-instructional, self-pacing packages using the PLATO System are employed for the twelve basic science disciplines (Sorlie and Jones, 1976). As Doull and Walaszek (in Votaw and Farquhar, 1978) inform us, a large system of programmed instruction named CATS (Computer Aided Teaching System) has been developed at the University of Kansas Medical Center, Department of Pharmacology. It consists of a large number of tutorial programs as well as an item bank with examination questions for testing. This system is now in use at fifty medical and pharmacy schools in the United States and Canada, (Votaw and Farguhar, 1978), including McGill. The Harvard Univerity Medical School teaches a number of subjects in the discipline of medicine by the use of Mumps Programming Language (Lefever and Johnson, 1976). The same language is used in the teaching of various subjects in pharmacology at Southern Illinois University, School of Medicine. Subjects in pharmacology are also taught at several universities using other systems.

At the University of California, San Diegeo, preclinical medical students are taught the fundamental skills of history taking, the framework of physical examination and some specific problem solving methods by a CAI system. The problem consists of an interactive encounter with model patients which critique the student gathering the clinical data elements, and a set of multiple choice questions on basic science and clinical topics. This approach is to test the feasibility of the usual in-hospital teaching of clinical skills that preceeds the clinical period (O'Neil, Sewall and Marchand, 1976).

Some attempts in medical education have concentrated on evaluating the effectiveness of the tutorial CAI with traditional lectures. Thus, in a study at the University of Oklahoma, two topics in Physiology were taught by either of the two methods, each in one session period. A month after the evaluation test, students were switched with respect to experimental treatment and material taught. No significant difference was found between the two groups in the test. Knowledge of material was recalled with equal effectiveness by both groups (Thies, Harless, Lucas and Jacobson, 1969). Meyer and Beaton (1974) compared the effectiveness of a CAI program with a human tutor on the same material in physiology; no difference was found in students' performance.

Major Applications of Problem Solving (Simulation) Techniques

Clinical problem solving in the form of patient simulation (i.e. patient management problems) has met broad acceptance in the field of medical education. The student, in the role of the physician, is expected to make a series of interdependent decisions throughout the various stages of the diagnostic and treatment work-up, from the consequences of which he gets feedback exactly as in the clinical solving situation in the hospital. The patient management problem is being used by the various medical education institutions in the form of written simulation, computer-based simulations and simulations with actors in the role of the patient. In all three forms the general approach is relatively similar, starting with the statement of the problem--as described by the patient or a relative--followed by a list of specific interventions (physical examination, lab tests, drug prescription etc.) or of general strategies. From these, the initial approach as well as subsequent approaches can be selected. As each decision is made by the problem solver, information regarding the results of it is provided in a realistic verbal or visual form (McGuire, Solomon and Forman, 1976). Clinical simulations were initially developed for evaluation purposes as dynamic tests of the clinical skill instead of the usual multiple choice questions. The pioneering work of McGuire (1963), McGuire and Babbot (1967), Williamson and McGuire (1968) and Barrows and Abrahamson (1964) in proposing, developing and using simulations, is of interest here, as well as the more recent studies on the reliability and validity of simulations as tests (Goran, Williamson and Gonnela, 1973; Palva and Korhonen, 1976).

The use of patient simulation as a systematic teaching device is a recent development in medical education, the applications of which are not always of the same form. Thus, some involve a simple simulated patient-physician encounter without any external aid to the student, while others provide a teaching section after the end of the problem solving process. Non-aided simulations are used as teaching and practice instruments when it is believed that external aid with the process of clinical problem solving is boring and destroys the spontaneity of the situation. Such simulations of patient encounters designed for advanced medical students and practising physicians are used at the Harvard Medical School as supplements to other ways of instruction and practice in problem solving (Lefever and Johnson, 1976).

Major Applications of the Combined Approach

An interesting system, which embodies the tutorial, inquiry and problem solving aspect of instruction is the "MENTOR" System proposed by Feurzeig (1970). Conversational interaction during the process of the student's attempt to solve the clinical diagnosis problem stated by the computer is the main characteristic of this program. In particular, the student types requests for the patient's history or physical examination or makes an assertion. The computer types a response consisting of a comment or an answer to a question. (Feurzeig, Munter, Swets and Breen, 1964). A similar approach is used in the University of California, Los Angeles, School of Medicine, as a supplement to the usual fourth year clerkship in Surgery. The student is presented with a clinical case and is allowed to enter requests from a selection list, in order to establish a diagnosis and decide about the treatment, while receiving continuous direction from the computer (Hammidi and Fonkalsrud, 1970).

A teaching device which utilizes external aid is the CASE (computer aided simulation of the clinical encounter) System proposed by W. Harless at the University of Illinois, Chicago, College of Medicine, which uses static patient models; that is, variables do not change as a result of action. The student is expected to make a diagnosis and prescribe a treatment on the basis of initial patient description and collection of information from the patient, physical examinations and lab test results. After the treatment is prescribed, the program defines those concepts the author believes to be critical for design and treatment and identifies deficiencies in the problem solving approach (Votaw, 1978). Recently, at the same University, a new approach for evaluating students' clinical problem solving skill has been developed, which might serve as an interesting teaching instrument. The approach introduced by L. Solomon, is called the Sequential Management Problem (SMP) and allows the student to generate--instead of choosing from a list--the requests for obtaining the data he needs. Further, at every step, the student gets corrective feedback, so that the chance for cumulative errors is less than in the patient management problem (Berner, Hamilton and Best, 1974)

Another significant approach is employed at the University of Illinois, Chicago, using the PLATO System. It is a computer-assisted encounter providing for practice in problem solving skill and clinical judgement (Sajid, Lipson and Telder, 1975). The tutorial approach in which questions are asked from the computer after the presentation of the material, is followed by the inquiry approach at which the student gathers and sorts the information he needs for his answers, again with the computer aid. That is, the computer when requested can indicate whether an answer is right or wrong, can help by providing a sequence of information according to the kind of trouble the student has, and can even give the answer. In this process, the student can ask his own questions, specific or more general to the topic and thus direct his own learning process (Bitzer, 1970). Finally, an attempt similar to those described above, which involves a non-computer device, is made at McMaster University. The 'problem box' employed there, involves a clinical problem protocol with related audiovisual material; at the end of the protocol new questions appear concerning information from many areas of human biology which lead the student into more understanding of the solution of the problem; finally a list of further relevant learning resources is presented to the student (Barrows and Mitchell, 1972).

Of the three categories of individualized instruction in medical education referred to above, the combined approach seems to be the most promising but at the same time the hardest to plan and apply, especially in the form of a computer program. A genuinely effective program of this kind would appear to require the use of artificial intelligence techniques (e.g. Brown and Burton, 1977). Such techniques are difficult to implement and require elaborate software systems that are currently not widely available.

General Comments

Medical schools make use of these techniques--formally or informally--without any systematic attempt to examine what is learned by whom. The selective lack of systematic studies upon the effectiveness of the various teaching devices is contrasted with the numerous descriptions and reports on applications of these devices, a fact that manifests some immaturity in the field of medical education. Clinical simulations are becoming more and more structurally complete instruments, and even their reliability and validity is carefully studied. However, this simply means that the development of instrumentation for education has surpassed educational development and tends to go out of control In fact, one can hardly find a consistent study in the literature, measuring the effectiveness of simulations as teaching instruments. Besides, the very few attempts have been unsuccessful in that the findings cannot be interpreted and the variables involved cannot be controlled. Sometimes, researchers report a positive attitude of the students towards a device of the above types. However, a positive attitude may have nothing to do with learning effectiveness, and it is this effectiveness which has to be measured in addition to the evidence about the favourable remarks from the students.

The problem seems to lie in that medical educators do not seem to be concerned systematically with detailed components of learning, namely, type of learning and parameters like retrieval, transfer, etc. An ongoing study by Maatsch and his Associates is one of the few examples of such specific attempts. This researcher is attempting to identify which variables are important for a particular instructional method in causing optimal recognition, recall, application, problem solving ability etc. Maatsch has often used patient simulations--not computer based--designed to facilitate instruction on the knowledge and skills associated with diagnosis and management. Some of his early findings are, that feedback must be furnished only at the time a mistake in a student's answer occurs in order to make a significant difference, and that there is no significant difference between overt and covert students' responses (Maatsch, 1974).

Ambitious proposals for developing 'intelligent' simulation programs may lead to great improvements over conventional tutorial programs. They may also lead to the development of reliable instruments for the evaluation of the learning act and of clinical competency. With such a sophisticated pairing of simulation to teach and simulations to evaluate performance, we might be able to know what components of the medical 'knowledge' are learned by a simulation, and more about how a physician solves a clinical problem. At present, however, we do not even know whether a simulation can teach at all, or if it can teach, under what conditions this happens. Do the teaching effects of a simulation conform to the overall findings about the teaching effects of discovery and problem solving--referred to at an earlier section? Such a question, which assumes the breaking down of the issue into smaller parts, seems to be suitable for the present stage, before more complex questions are ready to be investigated.

CHAPTER II

PURPOSE OF THE STUDY

This study is designed to investigate some aspects of the interactions between variables referred to in the preceding chapter. It will be concerned with the issue of how the two different teaching strategies referred to at the beginning of Chapter I influence performance on two relatively different kinds of knowledge in medical education: (a) basic factual knowledge of the kind usually taught in basic science courses; (b) clinical knowledge and clinical skill of the kind usually taught in clinical courses. In particular, the immediate effects of learning by means of a computer simulation (the problem solving strategy) will be compared with a computer based teaching strategy based on the more traditional teaching approach of guided ('expository') learning. Thus, a tutorial computer program will attempt to teach the same content as the simulation program by means of the presentation of multiple choice questions with feedback and tutoring dependent upon the student's responses. Both strategies will be tested on the basis of a single teaching session. The combined approach of guided problem solving is not involved in the study because of the difficulty of implementing such an approach on the computer.

The task assigned to the subjects (second and fourth year medical students) in this study is the management of a case in clinical pharmacology: a patient suffering from aspirin poisoning. The task of the management of this clinical case is involved in both experimental treatments. The programs will be desribed in detail in the next chapter.

The dependent variable in the study is students' performance in terms of 'basic' knowledge related to the topic of aspirin intoxication and 'clinical' knowledge related to the management of the particular case: e.g. generate the first steps of a treatment approach by choosing (or proposing) them from a given list of alternative steps. The above types of knowledge, basic and clinical are taught in the present study by means of the two programs instead of by lectures or clinical practice. The issue of previous knowledge, general or specific to the topic, will be discussed in a later section.

According to what has been found in previous studies on the two general approaches for various types of non-medical material (i.e. arithmetic, etc,) the tutorial approach is good for initial learning of subject matter while the problem solving approach is frequently better when retention and transfer are sought. In the present study, it seems reasonable to expect the tutorial program to be effective in teaching basic knowledge since the material is involved in the tutoring session. The question is whether it is effective at all in teaching clinical knowledge. This is a possibility since the whole program deals with the process of solving the particular problem and the student is involved in the various steps towards the solution. Additional help is provided from the program by the option given to the student to review the material on each step of the process of treatment. Obviously, one might learn from this tutorial program by merely recalling how the program was going through the successive steps towards the solution. Malin (1974) has found that this is possible, from her study which required retrieval from the memory of the sequence of learned formulas which formed a solution path to a problem.

On the other hand, the simulation program would not be expected to contribute considerably to initial basic subject-matter learning--something that has been asserted by the studies in the previous chapter. However, it is designed to enhance performance on the test of initial clinical knowledge. Most discussions of learning by problem solving emphasize the transfer effects rather than the initial learning. However, it seems reasonable to suppose that succesful initial learning may be harder to achieve because it demands a greater understanding of the specifics of the problem. Thus, in the present study, an attempt is made to evaluate whether initial clinical learning is at all possible under the problem solving approach.

The relative amount of learning by this approach is closely related to the amount of basic knowledge that the student has before entering the learning situation (unlike the tutorial approach which assumes a certain amount but can directly teach what is missing). In fact, the first point that theories of the problem solving approach emphasize--something that was elaborated earlier in the section on the nature of medical problem solving-- is the necessity of adequate knowledge of the basic material involved. On the other hand, clinical knowledge is also important as it should be expected to increase the probability of discovering a proper procedure for solving the problem.

For the above reasons, the variable of previous knowledge is also included in the study. Two levels of education are represented by two groups: second and fourth year medical students. Fourth year students had been taught the basic scientific subject matter (mainly related to the area of acid-base physiology). They had also, of course, received a much longer exposure to clinical training. It seems reasonable to expect that these students should perform better than the second year students on the simulation program. The effectiveness of the tutorial program, on the other hand, should not depend so much on the amount of previous knowledge. On the contrary, this latter program should tend to minimize the differences between the two groups.

On the basis of these considerations, it seems reasonable to ask the following questions:

(a) Is the tutorial program superior to the simulation program in teaching basic scientific knowledge?

(b) Is the simulation program superior to the tutorial program in teaching clinical knowledge?

(c) Is there an interaction between level of knowledge (2nd or 4th year) and the relative effectiveness of the two programs?

CHAPTER III

GENERAL METHOD

The present study involves two experiments, one with students in the fourth academic year and one with students in the second year. The two experiments, apart from involving two different samples of subjects also utilize slightly different experimental designs. The description of the common features will be treated in this chapter. The features which differ will be described in Chapters IV and V.

Apparatus

The hardware used in this study consisted of four Volker Craig 303 CRT display terminals. Three of these terminals, which belong to the Department of Pharmacology, were located in the medical building of McGill; one terminal was located at the Centre for Medical Education. All terminals were connected to the McGill IBM 370/158 time-sharing computer.

Computer-based Instructional Programs

The two programs used in the study treat the same topic of an emergency patient management case in clinical pharmacology, namely an aspirin intoxication case. A central feature in the material is acid- base physiology which underlies much of Internal Medicine and is a clinically useful topic. It is taught mainly in the courses in Physiology, Pharmacology and Internal Medicine. The two programs, though both interactive in nature, present the topic in quite a different mode. Both were written by Dr. J. Kreeft of the Department of Pharmacology, McGill University.

a) Tutorial Computer-based Program.

This program was coded in a CAI author language named CAN VI (developed by W. Olivier of the University of Toronto) in a version adapted to the McGill University System for Interactive Computing (MUSIC) by the Department of Educational Psychology and Sociology at McGill with the help of the McGill Computing Centre. Like most of the tutorial programs used in education, this program is frame-oriented. Each frame included one multiple choice question as well as comments and theoretical elaboration corresponding to each of the alternative answers of the question. Every time the student chooses an answer from the alternative answers of the multiple-choice question which appears at the beginning of each frame, he gets feedback about the correctness of his choice and more information about it. If he has answered correctly he is given the option of either proceeding to the next question-frame or seeing a review of the other possible choices with the corresponding comments and more information. The review follows automatically when the student has not answered correctly. The student is also given the option of going through a question again.

At the beginning of the program the student is presented with the clinical history of an adult who has ingested about 15 mgs of acetylsalisylic acid an hour before admission to the Medical Intensive Care Unit (MICU) of a hospital. Results of the physical examination of the patient during the first hour prior to admission also appear in the program. The questions that are addressed to the student concern the further examination of the patient, the diagnosis of his acid-base and electrolyte status and eventually his treatment. The program proceeds with the outcome of the initial treatment followed by a repetition of the whole clinical process (diagnosis and treatment) at 2, 4, 6, 8 and 15 hours post admission, when new values of the patient's condition are obtained. With the final diagnosis and treatment procedure the patient is in a condition that allows him to leave the MICU. Thus, each frame of the program is a step towards the solution of the problem of the management of this patient that proceeds in the way the expert knowledge of the author suggests and is independent of the student's prior responses.

The total number of frames within the program is 11; a sample of one frame as it appears on the screen is given in Appendix 1. For reasons of parallelism with the computer simulation program, described below, this tutorial CAI is non-personalized, which means it does not use an intimate conversational style, and does not use the student's name. Further, it does not grade the answers the student gives.

b) Computer-based Simulation Program.

This program, which is a clinical simulation of aspirin intoxication, has been available to McGill students since Winter 1975/76. It was initially written in BASIC and rewritten in Summer 1976 in FORTRAN to run under MUSIC. It is a complex program with many subroutines. One of these, ERSATZ, is of key importance as it contains a set of mathematical functions that relate time, treatment and various physiological compensatory mechanisms. It is claimed that this set of functions behaves with sufficient complexity to realistically simulate the effects of all the clinically important variables and their relationships (Kreeft, 1977). The program in this way allows the user-- who is playing the role of a physician placed in charge of the case at the MICU--to move into an open-ended problem field and take a series of sequential inter- dependent decisions at the various stages of therapy while the factor of time is also involved in a more or less realistic way.

At the beginning of the simulation, the clinical history and the patient's physiological changes during the hour prior to admission are presented in a way similar to that in the tutorial program. After reading the case history, the student is expected to manage the patient by entering requests from a list of therapeutic and diagnostic commands. A copy of the list with the alternative requests appears in Appendix 2.

The first decision point (i.e., when the student enters requests) always occurs one hour after ingestion of aspirin. Thereafter, the student determines the time to elapse until the next decision point by specifying that time to the computer. At each request the student can initiate, stop or modify treatment, ask for nursing findings, and ask for lab tests the results of which return in about 30 minutes. The student can also get a review of the current treatment or a chart review of the patient's input/output rate. When the patient suffers a cardiac arrest, the student faces an automatic decision point for which a cardiopulmonary treatment is available. The physiological changes resulting from a cardiac arrest are taken into account by the simulation. The termination of the simulation, which occurs after the patient is cured and can leave the MICU or when the patient dies, is followed by a treatment and chart review together with cost analysis. A sample of the format of the program as it appears on the screen is given in Appendix 3.

Tests

The state of knowledge of the students was tested before and after the experimental treatment by two similar--but not identical--tests, with material relevant to that involved in the programs. The two tests were constructed by Dr. J. Kreeft. Fifteen questions are included in each test, twelve of which are objective ones--in the format that appears in the examinaion of the National Board of Medical Examiners--and three are so-called 'sequence' questions. The questions of the latter type were designed for the diagnosis of problem-solving skills in patient management. They present a problem on aspirin intoxication. The student is asked to choose the steps he would follow from an extensive list or even to propose steps that are not included in the list. These questions, which are assumed to require cognitive skills more complex than simple recall, are scored both in terms of the importance of the selected steps and in terms of the sequence that is proposed. A copy of the two tests is given in Appendix 4. The questions of both tests were classified according to their content into a) 'basic' questions, measuring factual knowledge of the type that is taught in basic sciences during the first two years of medical education, and b) 'clinical' questions, measuring knowledge and skill of appropriate clinical procedures. The classification into the two categories was made by the author of the tests and was confirmed by another member of the Department of Pharmacology of McGill. In the one test (Test A) 'basic' and 'clinical' questions account for 32.5% and 67.5% of the total score respectively, while in the other test (Test B) the figure is 30.0% and 70.0% respectively. Criteria for scoring were developed in advance even for the unexpected answers in 'sequence' questions. A Kuder-Richardson Formula 20 Reliabilty coefficient of .77 was found for Test A and one of .81 for Test B. The construct validity for the two tests, expressed as the correlation between scores in the test administered before the experimental treatment and scores in the examinations of the National Board of Medical Examiners, is .56.

Questionnaire

As the time interval between the administration of the first test and the experimental treatment was 3 - 9 days, each subject was given a short questionnaire before exposure to the program. In the questionnaire it was asked whether the subject had learned anything related to aspirin intoxication since the administration of the first test--something that they had been asked to avoid doing. In case the subject had learned something relevant, he was excluded from the study. In the questionnaire it was also asked whether subjects were familiar with the simulation program which has been available in the Department of Pharmacoloy long before this study was carried through. A copy of the questionnaire is presented in Appendix 5.

Procedure

Letters describing the project were sent from the Centre for Medical Education to fourth and second year students of Medicine, McGill University, requesting their participation in the study. A few days after letters were mailed, students who had not already responded, were contacted by telephone. A copy of the letter addressed to second year students appears in Appendix 6.

These procedures resulted in a sample representing approximately one third of each of the two classes. The experiment required two sessions. In the first session--a few days before the experimental treatment--students had to take the first test which would estimate their knowledge on the topic of aspirin intoxication. The reason for this time lag between the two sessions was the elimination, or the diminishing, of carry-over or practice effects from the initial test over to the second test which was administered immediately after the treatment. Avoidance of carry-over effects was also the reason for developing two 'parallel' tests instead of one (pretest, post-test). The students of each class were randomly assigned to the two experimental treatments which were the two computer-based programs. The average time required for the tutorial program was 45 minutes and for the simulation 80 minutes. For the latter program, which is longer than the former, partly because it includes extensive instructions about its operation, the minimum amount of exposure time was 50 minutes and the maximum 110 minutes. During the treatment session, students were supervised and aided in matters such as computer operation, etc.

Treatment of Data

Scores in overall, 'basic' and 'clinical' performance from the two tests ('before' and 'after') were analysed in terms of the two treatments (tutorial and simulation program). Scores from 'clinical' questions were classified separately, according to objective clinical questions and 'sequence' clinical questions, for the inspection of possible differences arising from the different question format. As no such differences were found to exist, these analyses will not be reported and scores from both types of clinical questions will appear in one category--'clinical'. The total number of questions in each test amounts to 100 points, therefore, scores in 'overall' performance are percentages. Moreover, scores from 'basic' and 'clinical' questions separately, have been transformed so that they also appear as percentages.

The factor of previous experience with the simulation program has not been taken into consideration since the performance of the few students who were familiar with it was not found to be different from the group performance.
The approach to the analysis of data for the two experiments, involves presentation of tables with mean scores and standard deviations for every group, followed by analysis of variance and by multiple comparision between mean scores.

CHAPTER IV

EXPERIMENT I

Method

Subjects

Experiment I involves students from the fourth year of the Faculty the Faculty of Medicine at McGill University. The McGill medical curriculum is a four year program. The basic science courses occur primarily in the first two years. Pharmacology is taught from September through December of the second year. Clinical training begins in the middle of the second year. In January 1977, a revision of the medical curriculum introduced a three-month program of basic science after the end of the clinical work phase of the fourth year in the month of January. The aim of this innovation is the reconsideration of the basic science material in the light of the acquired clinical experience. Thus, at the time of this study, in March 1978, the fourth year students were attending classes on basic science options.

The sample in this experiment consisted primarily of students of high and average levels of achievement. The students ranged in age from 23 to 44 with the average age being 26.2. The majority of the students were males. A total of 3 students withdrew after the first test. In addition, incomplete data for another 3 students forced their exclusion from the sample. The data for these students was not included in any of the analyses. The number of students for whom complete data is available is given in Table 1.

Procedure

In this experiment, test A was administered to all students as the 'before' test (i.e., the pretest) and test B as the 'after' test (i.e., the post test). In this way, overall performance scores, as well as scores in 'basic' and 'clinical' questions, were analyzed in terms of two variables, treatment and 'before-after' administration of test.

Results and Discussion

The student-computer encounter did not give rise to any problems, even for students who were using the computer for the first time. Students were fairly happy with the programs, especially with the simulation. However, whenever the patient died (something that happened most of the time), frustration was obvious. The number of students who were already familiar with the simulation program is shown in Table 2. The percentage mean scores of these students in overall 'basic' and 'clinical' performance before and after the two treatments were calculated separately and were contrasted with the corresponding mean scores of the whole group. The table in Appendix 8 which involves this data indicates that no major differences seem to exist between the two sets of mean scores, with one noticeable exception. This is the mean score in clinical performance of the 'familiar' group before the simulation treatment which is considerably higher than that of the whole group (means of 67.4 vs 56.7). As, however, this difference ceases to exist after the exposure to the program (means of 45.0 vs 45.6) this fact can be taken as a

TABLE 1

TOTAL NUMBER OF FOURTH YEAR STUDENTS YIELDING COMPLETE DATA

Treatment	Original	Withdrawals	Missing Data	Final
Tutorial	31	3	1	27
Simulation	29	0	2	27
Total				54

TABLE 2

NUMBER OF FOURTH YEAR STUDENTS FAMILIAR WITH THE SIMULATION PROGRAM

	Last time of contact		
Treatment	More than one year ago	Less than one year ago	Total
Tutorial	5	1	6
Simulation	3	1	4
Total			10

coincidence because of the following contradiction: if the 'before' superiority in performance is due to familiarity with the simulation program, how can it be lost within one more single exposure to the program? Besides, the time gap since the last time of exposure which is for most of the 'familiar' students more than one year, does not justify such a strong effect of the program. What can be assumed here rather, is that differences in students' initial performance tend to diminish after exposure to simulation. This statement will be elaborated more at a later point of the thesis. Concluding that the above mentioned difference is not a major one, the individual scores of the 'familiar' students are treated together with the scores of the remainder of the group.

Table 3 presents percentage mean scores and standard deviations of individual performance scores (overall, basic and clinical) in the two tests before and after the two treatments for 54 students. On inspection, the existence of an interesting trend in the pattern of mean scores becomes evident, namely that performance after the exposure to the tutorial program improved considerably while performance after exposure to the simulation remained the same and even decreased. In an attempt for a more precise statistical analysis the data was then submitted to a two-factor analysis of variance (treatment x before-after administration) for each of the three types of performance (overall, basic and clinical). A summary in terms of F values of the three tables of analysis of variance which appear in Appendix 7, is given in Table 4.

As the data in Table 4 reveals, the two treatments

TABLE 3

% MEANS AND STANDARD DEVIATIONS OF INDIVIDUAL PERFORMANCE SCORES BEFORE (TEST A) AND AFTER (TEST B) THE TWO TREATMENTS.

		BEFORE (Test A)		AFTER (Test B)	
Program	Performance in	Mean	SD	Mean	SD
	Overall	50.0	7.5	62.5	11.3
Tutorial (N=27)	Basic	45.3	19.9	70.4	15.6
	Clinical	52.3	11.7	59.0	13.9
	Overall	53.2	9.3	45.8	9.7
Simulation (N=27)	Basic	45.8	11.5	46.0	16.6
	Clinical	56.7	13.5	45.6	10.9

TABLE 4

SUMMARY OF THREE ANALYSIS OF VARIANCE TABLES INVOLVING SCORES IN 'BASIC' 'CLINICAL' AND OVERALL PERFORMANCE BEFORE AND AFTER THE TWO TREATMENTS.

	F		
Source of Variation	Overall	Basic	Clinical
Main Effects			
Treatment	12.87**	14.60**	3.47
Before-after	1.71	16.37**	.78
Interaction			
Treat x Before-after	30.13**	16.04**	13.50**

* $p < .05$

** $p < .01$

affect 'basic' performance in a significantly different way (F=14.6), and this also applies to overall performance. However, in 'clinical' performance, this significant difference between treatments is absent(F=3.47). Further, the factor 'before-after' reveals a significant difference only with respect to 'basic' performance(F=16.37). Interestingly enough, though, the interaction between treatment and 'beforafter' is significant for all three types of performance, thus, establishing that performance in test A (before) and B (after) varies according to the two instructional methods. Since analysis of variance as it is being used in the present study does not determine which of the pairs of sample means differ significantly, a multiple comparison analysis on an a posteriori basis is performed using the Tukey test and the Scheffe test.

Tables 5 and 6 demonstrate the outcomes of the Tukey (T) and Scheffe (S) comparison tests for the various pairs of means. The .05 and the .01 level of significance have been chosen for the two tests respectively, while differences below these levels are indicated by asterisks (*=p<.05, **=p<.01). As shown in the two tables, while students under both experimental treatments start with equal overall, 'basic' and 'clinical' performance, as measured by means of the 'before' test, they end up with significantly different performance from the two treatments on the "after" test. In fact, inspection of the mean scores of table 3 reveals that it is the tutorial program which is superior to the simulation program, for both 'basic' and 'clinical' material. As far as the simulation program is concerned, scores in 'after' test

TABLE 5

MULTIPLE COMPARISON BETWEEN MEANS OF TEST A (BEFORE) AND TEST B (AFTER) USING THE TUKEY-METHOD (T) AND THE SCHEFFE-METHOD (S)

Program	Performance in	$T_{(.05)}$	$S_{(.01)}$
Tutorial	Overall	*	**
	Basic	*	**
	Clinical	-	-
Simulation	Overall	*	**
	Basic	-	-
	Clinical	*	**

* $p < .05$

** $p < .01$

TABLE 6

MULTIPLE COMPARISON BETWEEN MEANS IN TUTORIAL AND SIMULATION PROGRAMS FOR TEST A AND TEST B USING THE TUKEY-METHOD (T) AND THE SCHEFFE-METHOD (S)

	BEFORE (Test A) Tutorial-Simulation		AFTER (Test B) Tutorial-Simulation	
	$T_{(.05)}$	$S_{(.01)}$	$T_{(.05)}$	$S_{(.01)}$
Overall	-	-	*	**
Basic	-	-	*	**
Clinical	-	-	*	**

* $p < .05$

** $p < .01$

are significantly lower than those in the tutorial program in both 'basic' and 'clinical' material. Further evidence for this is given in table 5 where 'before' and 'after' test are compared in terms of each treatment. For the tutorial program, the 'basic' performance after the treatment is significantly higher than before treatment, while 'clinical' performance does not reveal such a difference. On the contrary, under the simulation program, it is only 'clinical' performance which reveals a significant difference. Unfortunately, it is a decrease rather than an increase.

At this stage of the analysis of data, however, some reservations should be kept, on account of the fact that the two tests A and B have not been compared in terms of relative difficulty. In fact, a systematic consideration of this aspect will take place in Experiment II, where the two tests are counterbalanced. Nevertheless, from the mere comparison of performance in the 'after' test (test B) in table 6 between the two groups, it becomes apparent that the tutorial program is indeed, significantly superior in all three types of performance. A more extensive discussion, though, will be more meaningful after the results from Experiment II are analysed.

In an attempt towards a further analysis of the present set data, a comparison was performed between students who indicated high and low performance before the experimental treatment. Using the 25th percentile as a score criterion, 14 students (7 from each treatment group) were chosen for the group of high performance and 14 students for the group of low performance, on the basis of their overall score in test A (before treatment). Table 7 illustrates the mean scores of the

TABLE 7

FIRST (HIGH PERFORMANCE) AND FOURTH (LOW PERFORMANCE) QUARTILE MEAN SCORES OF FOURTH YEAR STUDENTS BEFORE AND AFTER THE TWO TREATMENTS.

		Tutorial			Simulation		
		HIGH (N=7)	whole group	LOW (N=7)	HIGH (N=7)	whole group	LOW (N=7)
BEFORE	Overall	58.7	50.0	41.0	64.0	53.2	40.2
	Basic	52.6	45.3	40.6	50.4	45.8	39.5
	Clinical	61.3	52.3	41.1	70.4	56.7	42.9
AFTER	Overall	63.0	62.5	67.7	48.1	45.8	37.3
	Basic	73.0	70.4	73.6	42.7	46.0	38.1
	Clinical	60.8	59.0	65.1	50.4	45.6	36.9

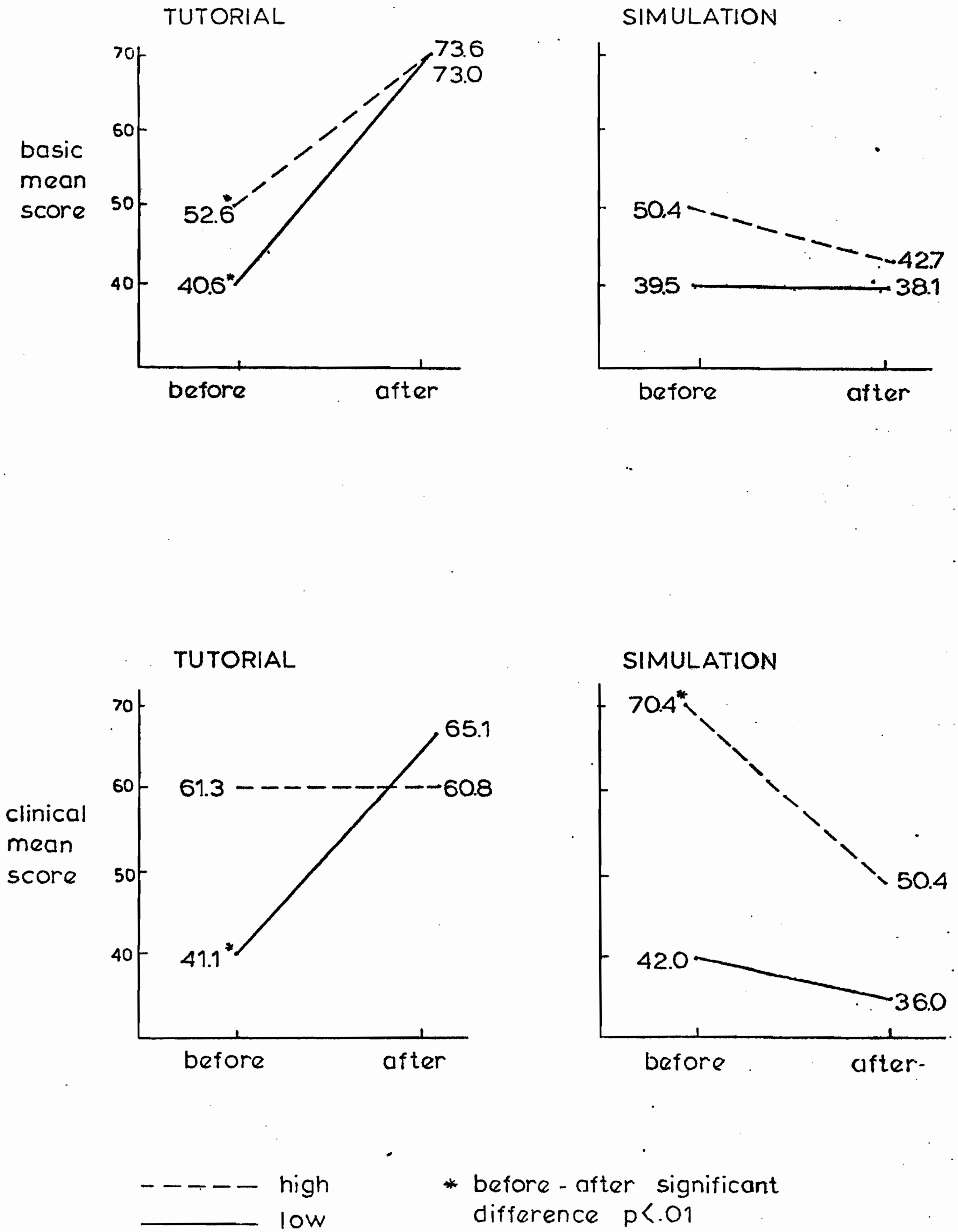
four groups in terms of overall, 'basic' and 'clinical' performance, both before and after the two treatments. The mean scores of the whole group are also included for comparison purposes.

As this table reveals, with respect to the whole group (average) performance, high performance students started with a superiority in both 'basic' and 'clinical' material. After both treatments though, the high performance group scored almost the same as average in both types of material. Low performance students, on the other hand, starting with an inferiority mainly in 'clinical' material managed to even exceed average (whole group) performance after the tutorial program. After the simulation program, however, low performance students remained low with respect to the average performance As is apparentin the comparison so far, the two treatments do not affect students of high and low performance in the same way, since high performance students benefit much less from the tutorial program and lose more under the simulation, compared to the low performance students. The picture will become more complete after inspection on the direct relationship between the two groups. This relationship is depicted graphically in Figure 1.

The top two graphs of Figure 1 illustrate the mean score of the high and low performance group in 'basic' material before and after the two treatments. The mean score of the two groups in 'clinical' material is demonstrated in the bottom two graphs. Asterisks indicate significant differences in before-after pairs when a t-test is used. Interestingly enough, the pattern of the before-after performance is

FIGURE 1

COMPARISON BETWEEN FOURTH YEAR STUDENTS WITH HIGH AND LOW PRETEST SCORES



different for the two groups and often interacts as in the case of clinical performance under the tutorial, while a similar pattern appears in clinical performance under the simulation treatment. A remarkable feature in these two cases that signifies an interaction, is that while the 'before-after' performance of one of the two groups provides a more or less horizontal line (no difference between means) the performance of the other group has changed significantly from the 'before' to the 'after' test. This trend is not so obvious when performance in basic material is concerned.

In general, it becomes apparent that low performance students benefit considerably more than high performance students from the tutorial treatment and loose considerably less (if at all) than the high performance students under the simulation treatment. Although this trend is mostly evident for the former treatment, data under the simulation is also interesting, especially in 'clinical' material where the difference between the means of the 'high' and 'low' group, from 28.4 points went down to 14.4 points. It should be pointed out that in all cases except 'clinical' under simulation, the scores of the two groups before the treatments differ singificantly at least at the .05 level when a t-test is used. After the treatments, though, no such difference remains.

CHAPTER V

EXPERIMENT II

Method

Subjects:

Experiment II involved students from the second year of Medicine at McGill. In April 1978 when this study was carried out, students were attending the last classes in basic science courses while their involvement in clinical work in the hospitals had started three months earlier.

The students ranged in age from 21 to 32 with the average age being 23.8. The majority of the students were males. A total of four students withdrew after the first test administration, and one student at the time of the experimental treatment. Further, one student, having reported that he had studied on the topic after the first test administration, was excluded from the sample. The final number of students yielding complete data--which was all that was used in the study--is given in Table 8a.

Procedure

As has been mentioned earlier, the experimental design for the second year students is not the same as that for the fourth year students. Here, counter-balancing of tests A and B was introduced in order that any possible discrepancies in the nature of the two tests be controlled; in this way half the students took test A before treatment and test B after, and the other half took them in the opposite order. This

TABLE 8a

TOTAL NUMBER OF SECOND YEAR STUDENTS YIELDING COMPLETE DATA

Treatment	Original	Withdrawals	Missing Data	Study after first test	Final
Tutorial	28	2	1	1	24
Simulation	26	3			23
Total					47

TABLE 8b

THE COUNTERBALANCED DESIGN OF EXPERIMENT II

GROUP	N	CONDITION	PRE TEST	POST TEST
G1	10	Tutorial	A	B
G2	14	Tutorial	B	A
G3	12	Simulation	A	B
G4	11	Simulation	B	A

could not be done for the fourth year students in Experiment I because of technical problems with the computer system. Figure 8b illustrates the experimental design applied for second year students as well as the number of students in each group.

According to this procedure, percentage mean scores for overall, basic and clinical material were obtained with respect to three independent variables: treatment, before-after test administration and name of test (A,B).

Results and Discussion

The student-computer encounter did not give rise to any problems. Students were happier with the tutorial program than with the simulation, in which the patient died most of the time. The number of students who were already familiar with the simulation program, appears in Table 9. All students had the last contact with the program less than one year ago. As this number was very small and, on inspection, scores did not differ from the scores of the whole sample, these students were included in the groups. The percentage mean scores and standard deviations of the 4 groups, before and after the two treatments, appear in Table 10.

A primary concern on inspection of this table is to determine whether the two tests A and B are different. Indeed, mean scores in the two tests before treatment seem to be similar only as far as the basic items are concerned. Thus, the mean obtained in test A is 46.1 for the group under tutorial treatment and 48.0 for the group under simulation, while the means obtained in test B are 49.4 and 48.5 for the

TABLE 9

NUMBER OF SECOND YEAR STUDENTS FAMILIAR WITH THE SIMULATION PROGRAM

		Last time of contact	
Treatment	Order of Test	More than one year ago	Less than one year ago
Tutorial	A-B	-	1
	B-A	-	3
Simulation	A-B	-	2
	B-A	-	1
Total			7

TABLE 10

MEANS AND STANDARD DEVIATIONS OF INDIVIDUAL SCORES BEFORE AND AFTER THE TWO TREATMENTS FOR ORDER OF TESTS A-B AND B-A.

	Sequence of Tests		BEFORE MEAN	BEFORE SD	AFTER MEAN	AFTER SD
TUTORIAL	G1 A-B	Overall	47.1	13.7	59.9	12.6
		Basic	46.1	13.6	65.8	20.6
		Clinical	47.5	16.6	57.3	10.8
	G2 B-A	Overall	43.4	9.4	59.8	8.5
		Basic	49.4	14.0	58.2	14.0
		Clinical	40.8	11.9	60.5	13.3
SIMULATION	G3 A-B	Overall	53.1	12.8	42.8	8.7
		Basic	48.0	19.7	47.2	17.8
		Clinical	55.5	15.3	40.9	11.9
	G4 B-A	Overall	43.3	8.5	49.7	11.8
		Basic	48.5	17.8	49.6	19.3
		Clinical	41.0	7.1	49.7	13.1

two groups respectively. The clinical performance before treatment manifests a difference between the two tests significant at the .05 level when a t-test is used. In view of this finding, a further interpretation of Table 10 with respect to the mean scores after treatment would not be very helpful at this time. Nonetheless, apart from this difference in the level of difficulty of the clinical material, the counter-balancing of the two tests does not manifest serious disturbances in the picture as no other major difference seems to exist between the two tests.

As Tables 11 and 12 illustrate, when t-tests are performed on various pairs of groups, the two tests (A and B) provide fairly parallel statistical features in most instances. In Table 11, the mean score of test A before treatment is compared with the mean in test A after treatment for each of the two treatment groups, and the same is done for test B. As can be seen, the levels of significance for the difference of each pair of means provided by the two tests are fairly similar. An interesting feature in this table is a significant difference in clinical performance before and after the tutorial treatment which, actually, implies improvement after treatment. In Table 12, comparisons are performed for paired groups that is, between the mean scores that the same students have provided before and after treatments. Here again, clinical performance reveals a significant difference under the 'simulation' treatment; unfortunately though, this is a misleading coincidence because while pair 'B before - A after' (p= .019) indicates a positive treatment effect, pair 'A before - B after' (p= .020)

TABLE 11

SIGNIFICANCE LEVELS OF T-TESTS BETWEEN 'BEFORE' AND 'AFTER' TREATMENT PERFORMANCE FOR TEST A AND TEST B UNDER THE TWO TREATMENTS.

	TUTORIAL		SIMULATION	
	TEST A (N=20)	TEST B (N=20)	TEST A (N=22)	TEST B (N=22)
	BEFORE-AFTER	BEFORE-AFTER	BEFORE-AFTER	BEFORE-AFTER
Overall	.011	.008	.882	.962
Basic	.084	.062	1.000	.923
Clinical	.026	.005	.442	.988

TABLE 12

SIGNIFICANCE LEVELS OF T-TESTS FOR PAIRED GROUPS BETWEEN TEST A AND B IN BOTH SEQUENCES OF TEST ADMINISTRATION AND FOR BOTH TREATMENTS.

	TUTORIAL		SIMULATION	
	(N=10) A BEFORE-B AFTER	(N=14) B BEFORE-A AFTER	(N=12) A BEFORE-B AFTER	(N=11) B BEFORE-A AFTER
Overall	.003	.006	.087	.104
Basic	.010	.119	.902	.818
Clinical	.030	.003	.019	.020

signifies a negative treatment effect. In view of this inconsistency, and having sufficient evidence about the similarity of the two tests apart from the level of difficulty, it seems reasonable to merge the scores, thus keeping only two independent variables: treatment and 'before-after' test administration. Such a merging, of course, assumes the risk of an increased variance especially in clinical performance data. The advantage though of a more concise statistical picture with a larger number of students per group and free from contradictions of the type that were encountered in the clinical data, makes the attempt worthwhile.

Table 13 which is a summary of three two-factor analyses of variance (the complete tables are given in Appendix 9) for the three types of material, offers a clear picture of the data. As this Table illustrates, the main effects for treatment in 'basic' ($F=2.86$) and 'clinical' ($F=2.17$) performance could not approach conventional levels of significance as they did for the overall performance ($F=4.19$). This trend is related to the fact that the variable 'before-after' has provided significant differences for all three types of performance--thus neutralizing the factor of treatment. Finally, 'basic' performance ($F=3.53$), unlike 'clinical' ($F=11.46$) and overall ($F=13.89$) fails to evidence a significant interaction of treatment by 'before-after'. For a more detailed view of the particular features that account for the above observations, each pair of groups was compared using the Tukey and the Scheffe tests for multiple comparisons. Tables 14 and 15 illustrate the significant differences

TABLE 13

SUMMARY OF THREE ANALYSIS OF VARIANCE TABLES INVOLVING MEAN SCORES IN 'BASIC', 'CLINICAL' AND OVERALL PERFORMANCE BEFORE AND AFTER THE TWO TREATMENTS.

Source of Variation	Overall	Basic	Clinical
Main Effects			
Treatment	4.19*	2.86	2.17
Before-After	7.44**	3.66*	4.59*
Treat x Before-After	13.89**	3.53	11.46**

* $p < .05$

** $p < .01$

between mean scores when the level of significance chosen is p= .05 (*) for the Tukey test and p= .01 (**) for the Schefe test. In Table 14 which involves comparisons between means obtained before and after each treatment, the mean scores and standard deviation of the merged data are also included. In fact, as the standard deviations reveal, the variance after the merging of the two tests (A and B) does not increase to unsatisfactory levels, something that makes results reasonably reliable.

Table 15 involves comparisons between the means of the two treatment groups that were obtained before exposure to the program and also after the exposure. Evidently, no significant difference appears in the table before the program exposure, whereas such a difference becomes apparent after the exposure.

From both tables it becomes evident that the significant differences within the factor 'before-after', revealed in the analysis of variance, are almost exclusively due to a significant improvement of performance after the 'tutorial' treatment whereas simulation in itself, does not provide any such difference (Table 14). As this difference due to the tutorial program does not exist at the 'before' test (Table 15), the difference between the two treatments could not appear to be significant in the analysis of variance. The significant interaction between treatment and 'before-after' in 'clinical' performance, supports this statement while the lack of such a significant interaction in 'basic' performance is only due to less extreme values, since the trend is the same with 'clinical'.

TABLE 14

MEANS AND STANDARD DEVIATIONS BEFORE AND AFTER THE TWO TREATMENTS AND MULTIPLE COMPARISON BETWEEN MEANS, USING THE TUKEY-METHOD (T) AND THE SCHEFFE-METHOD (S)

TREATMENT	BEFORE: Performance in	BEFORE: Mean	BEFORE: SD.	AFTER: Mean	AFTER: SD	COMPARISON: T(.05)	COMPARISON: S(.01)
TUTORIAL	Overall	44.9	11.2	59.8	10.1	*	**
	Basic	48.0	13.6	61.4	17.0	*	**
	Clinical	44.6	14.1	60.0	12.2	*	**
SIMULATION	Overall	48.4	11.9	46.1	10.7	-	-
	Basic	48.3	18.4	48.4	18.1	-	-
	Clinical	49.3	13.6	45.9	13.0	-	-

TABLE 15

MULTIPLE COMPARISON BETWEEN MEANS IN TUTORIAL AND SIMULATION PROGRAMS, BEFORE AND AFTER TREATMENT, USING THE TUKEY-METHOD (T) AND THE SCHEFFE-METHOD (S)

	BEFORE: Tutorial-Simulation T(.05)	BEFORE: Tutorial-Simulation S(.01)	AFTER: Tutorial-Simulation T(.05)	AFTER: Tutorial-Simulation S(.01)
Overall	-	-	*	**
Basic	-	-	*	-
Clinical	-	-	*	**

As the above results indicate, a superiority of the tutorial program over the simulation which is expressed to an equal extent for both 'basic' and 'clinical' type of material, is obvious for students from the second year.

CHAPTER VI

GENERAL DISCUSSION

A Comparison of the Results of Both Experiments

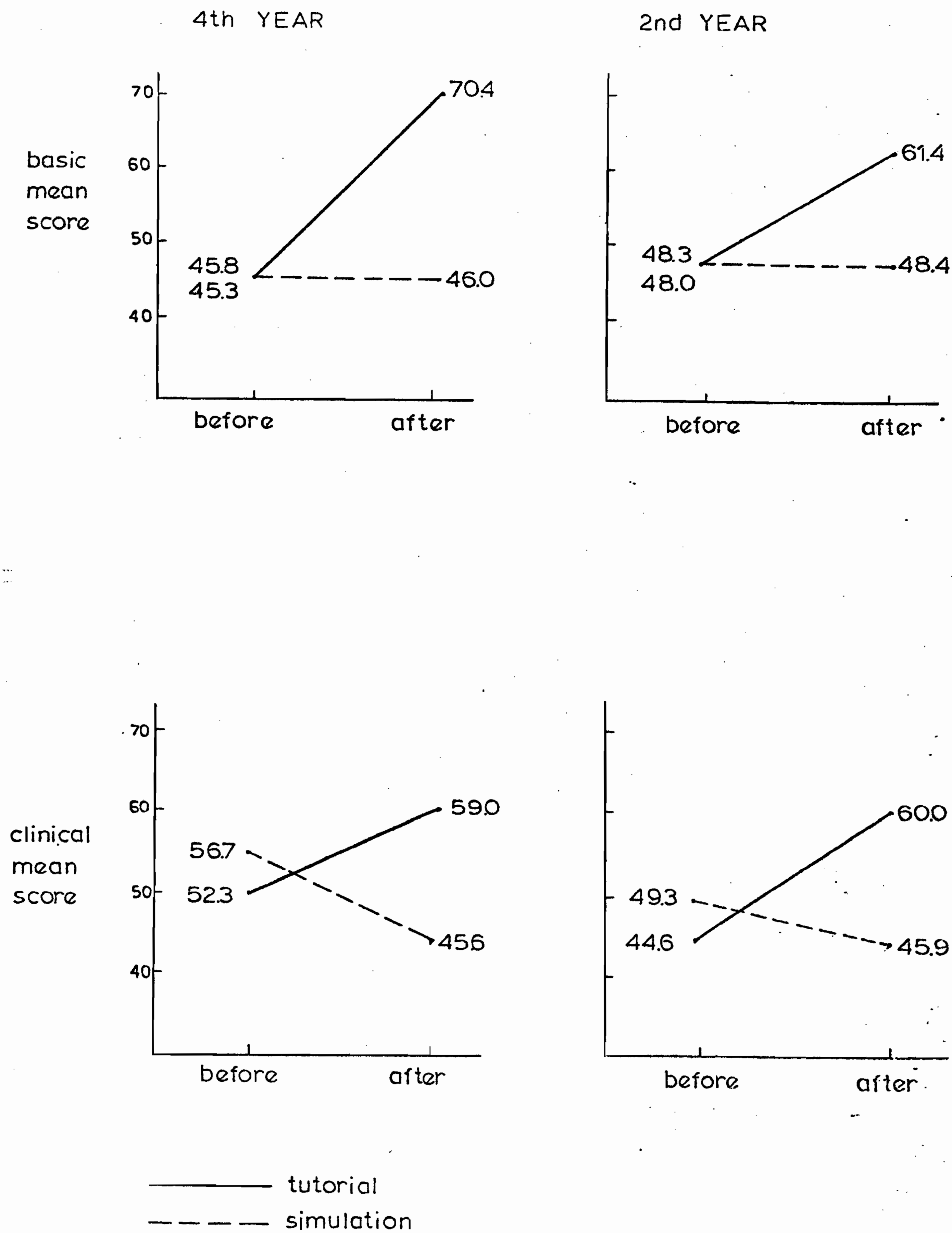
Part of the purpose of this study, stated earlier, refers to the relationship between the results from the second and fourth year medical students. Figure 2 depicts graphically a summary of these findings from the two classes. The data in this figure refer to the mean scores obtained by the two academic year groups before and after each experimental treatment for 'basic' and 'clinical' material. They have already been presented separately in Chapters V and VI. They are presented together in Figure 2 for convenience of comparision.

The interactions that the two treatments manifest are very similar for the two classes. The graph for basic material reveals the same trend of effects of the two treatments, namely, no change whatsoever, after exposure to the simulation treatment, while for the tutorial treatment, a considerable improvement appears which is especially obvious for the fourth year subjects. Another interesting feature here is that the two class groups do not differ greatly in their mean scores on basic material in the pretest (means of 45.5 for the fourth year group versus 48 for the second year group), indicating that both began with the same amount of 'basic knowledge'. Yet, the fourth year group, as was shown, benefitted more than the other from the tutorial treatment.

The comparative findings arising from the 'basic'

FIGURE 2

SUMMARY OF DIFFERENCES BETWEEN SECOND AND FOURTH YEAR STUDENTS.



material can be considered as fairly reliable, despite the fact that tests A and B were used in a different way in the two experiments. This is because no difference in the degree of difficulty was found between them with respect to basic questions, when the second year mean scores from the two tests were compared.

Coming now to the data on the 'clinical' material, the figure reveals the same pattern of interaction for both classes, namely, an overall positive effect of the tutorial program and a neutral-negative effect of the simulation program. Here, however, the amount of change after the treatments is not the same for the two classes, as the fourth year class shows no significant improvement after the tutorial program while after the simulation it shows a significant decline. On the other hand, the second year group reveals a significant improvement after the tutorial program and no difference after the simulation. This result gives rise to the conclusion that the simulation program had a negative effect for the fourth year subjects and no effect for the second year subject. Moreover, the present comparison reveals that the mean scores obtained by the two classes in 'clinical' material before treatment, are not as similar as they were for the 'basic' material; in fact, fourth year group has somehow provided higher means than second year. At this point, however, the difference in the difficulty of the two tests A and B with respect to clinical questions has to be taken into consideration. As will be recalled, the mean scores before treatment of the fourth year are the outcome of the test A administration while the analogous mean scores of the second year come from the administration of both A and B tests. However, test B was found to be more difficult than A and, therefore, to provide lower scores. Hence, the mean scores obtained by the second year (administration of both tests) must be lower than that of fourth year, because of the nature of the tests and not because second year began with a lower clinical knowledge. Interestingly enough, this implies that the two classes can be assumed not to differ in terms of 'clinical' knowledge before their exposure to the programs. By the same token, the mean scores after treatment for the fourth year subjects, which are as high as those of the second year, were obtained by administration of test B only. If these had come from both tests, as is the case or the second year mean scores, they should be higher; hence, fourth year 'clinical' mean scores after treatment can be assumed to be higher than the second year scores. According to this way of reasoning (i.e., a hypothetical assumption that counter-balancing of tests was done for both), the two classes reaveal the same pattern between 'before treatment' and 'after treatment' performance in 'clinical' material, namely, a considerable improvement after the tutorial program and no change after the simulation.

Overall, the graph that sets the two experimental outcomes side by side indicates the surprising phenomenon of no difference in the knowledge of the two different classes before the experimental treatments. Moreover, the effects of the two treatments appear to be the same for both classes, namely, a positive effect of the tutorial treatment on both 'basic' and 'clinical' performance and no effect of the simulation treatment on either type of knowledge. The fact that the fourth year subjects benefitted even more than the second year ones from the tutorial in 'basic' material and perhaps to some extent in 'clinical' arising only from the harder test B--is unlikely to be attributed to chance. Therefore, it deserves some attempt at interpretation.

Lack of Difference Between the Two Classes, Pre-treatment Knowledge

The major interest of this study centered on the effectiveness of the two instructional strategies. One variable though, of considerable importance was the different background of medical students which gave rise to the two experiments. Apparently, the two instructional strategies were expected to depend partially on the knowledge that the subjects possessed before treatment. Quite surprisingly, the students from the second year were found to possess before treatment an amount of basic and clinical knowledge equal to that of the fourth year on the topic of aspirin intoxication. As far as basic knowledge is concerned, fourth year students had most probably been taught the theory relevant to the topic for a second time in the Basic Science Options program, unlike second year students, and therefore should perform better. Further, with respect to clinical knowledge, these students should have provided higher scores before treatment if not because of experience in the relatively rare case of aspirin intoxication at least as an outcome of their clinical experience which could be transferred to the particular case.

Nothing like this happened though, and one has to look for an explanation. A basis for such an explanation with respect to clinical performance only, is available in the relevant literature. Palva and Korhonen (1976) compared the clinical knowledge of 3 groups of medical students at early stages of the clinical workshop period, at the beginning of the clinical clerkship period in internal medicine and at the end of the clerkship period, on a simulated clinical problem (drug-induced agranulocytosis). It was found that the total performance of the two first groups of students did not differ significantly, while the performance of the last group was significantly superior and with the same percentage of success as in the real medical practice. As the authors suggest, the gain in clinical knowledge is most rapid during the early phases of clinical teaching and does not improve greatly until the very late stage of clerkship. Further, the level of relevant knowledge of topics in internal medicine may be even lower at the beginning of the clerkship period than at the early stages of clinical work--mainly because of interference. In fact, the diagnostic skill of the subjects at the beginning of the clerkship period had remained the same as in the early clinical period, while their alertness to the therapeutic approach had even deteriorated. On the basis of this data, the two first groups of which correspond almost exactly to those of the present study, one might generalize and attribute the non-superiority in the fourth year clinical performance to the same factor, namely, exposure to too much information and some resulting confusion. Apparently here, second and fourth year students are similar to De Groot's weaker chess players, who lack a good organization of the repertory of their experiences. Thus, the phenomenon of interference might replace transfer of learning. transfer of learning.

In conclusion, it can be assumed that overall, the two classes did not differ in their pretreatment performance in terms of either basic or clinical material and it is on this basis that the discussion on the two teaching strategies must be carried out. Of course, the possibility that the tests used in the study were not sensitive to general clinical knowledge should be taken into consideration. Yet, Palva's and Korhonen's study with the evidence of no difference between the two first groups keeps the above possibility small. There is one detail, however, which indicates that clinical experience of the fourth year group has not remained totally latent or ineffective. The present study has provided some evidence about a positive attitude which fourth year students showed toward the simulation program unlike the second year students who seemed to prefer the tutorial program much more. Palva (1974) reports the same phenomenon in a study involving the measurement of clinical problem solving by means of clinical simulation tests. According to his contentions, the junior students found themselves under pressure when confronted for the first time with the problems while the senior students found the situation relevant and enjoyed it, regardless of the results.

The Effects of the Tutorial Program

The tutorial program gave rise to a significantly superior performance in basic material for both class groups. However, although it would be expected that the two groups should end up with levels of performance similar to each other (an expectation that became even stronger when it was found that the two groups began at the same amount of knowledge), this was not the case. The fourth year group showed a much better performance in basic material. In this case, the only variable whch was not the same for the two groups and therefore should be suspected, is the different background. This could be due to two possible reasons: a) fourth year students having been taught twice about relevant topics possessed some latent knowledge, which mainifested itself during the tutorial program and enhanced recall; b) fourth year students having had two more years of experience could use the information in the tutorial in a more constructive way. Both possibilities seem reasonable. Yet, it is not clear whether it was the particular computer based strategy which caused this difference in performance or if other teaching strategies could give the same outcome.

In general, though, the hypothesis about a satisfactory teaching outcome of the tutorial strategey on the initial learning of basic medical material was strongly confirmed. The success of the tutorial approach on 'basic' knowledge coincides with similar previous findings on basic non-medical material. However, a major contributor to the effectiveness of this approach that should not be underestimated is the familiarity students have with the deductive teaching methods. To the students' familiarity one might attribute, in part, the remarkable difference between the time needed for the exposure to the two programs (the tutorial program required much less time).

As far as clinical knowledge is concerned, this was also found to be positively affected by the tutorial program; second year students revealed a really significant improvement after the treatment, as was previously observed to be the case for the fourth year students. In fact, the effect for the fourth year students seems to be even stronger than that for the second year but, unfortunately, the evidence is not sufficiently statistically significant to deserve extensive discussion.

Thus, the tutorial program was found effective even with respect to clinical knowledge. According to general assertions on the role of tutoring and guided learning, no particular positive effect upon skill in problem solving has been reported. However, the point made by these studies is that tutoring cannot guarantee broad transfer which is necessary for skill in general problem solving. When a specific problem solving skill is concerned as in the present study, a minimum amount of transfer is required. Specific aspects of clinical knowledge and problem solving skills can be taught directly within each frame of the program, by feedback to the multiple choice questions and by tutoring. This, of course, will be learned in the form of rules with little flexibility and little possibility of transfer.

In the same form of learning, skill in planning the global steps to the solution of this problem may also be acquired. As mentioned earlier, with reference to the study by Johnson, Maler and Bass, the planning of the solution of a clinical problem requires the application of general strategies--in the form of rules--which are retrieved from long term memory and may be familiar to any physician (i.e., interpretation of the information in the problem, early generation of diagnostic hypotheses, etc.). In the particular problem solving approach here, these strategies are involved too. The tutorial program implied such global strategies without teaching them directly as it proceeded from frame to frame. Therefore, if a student did not know them before the treatment, he might have picked them up during it. All this is very likely to make a student relatively capable to solve a clinical problem on the same topic. Very possibly, some aspects of the problem solving process might have not been included in the program, and therefore not learned. Overall, however, the tutorial program is found capable to provide a significantly superior performance on subject-matter clinical material and also, as the findings reveal, superior than what the simulation can give.

Finally, it should be pointed out that there is another way of interpreting the findings on the effectiveness of the tutorial program on clinical questions. This possibility, which does not necessarily exclude the former explanation, is that students were able to show improvement with respect to clinical knowledge after the treatment by using their basic knowledge which had improved during the treatment. Finally, it is not unlikely that transfer of their general problem solving skill has also played a role.

The Effects of the Simulation Program

The main question regarding the simulation or problem solving approach was whether this would improve clinical knowledge and whether the amount of improvement would depend on the amount of knowledge one begins with. Basic knowledge was not expected to improve. This expectation about basic knowledge, which is based on data from general studies mentioned earlier, is perfectly well supported from the results of this study. In fact, the mean scores after simulation are almost the same as the scores before the treatment for both the fourth and second year groups. This is reasonable if it is assumed that the possible discoveries that may occur during the process primarily concern skill in problem solving. In this respect, basic knowledge is only part of the input to the problem solving process.

With respect to clinical knowledge, the simulation program did not yield any improvement but on the contrary gave a slight evidence of a negative effect. The expression 'slight evidence' is used here because the significanctly negative effect that was observed with the fourth year group is suspect due to the two different tests that were used. The evidence of a negative effect is attributed to confusion that inhibited learning during the exposure to the program. In this way, even the more experienced fourth year students did not do better than the second year students. The confusion appears to coincide with their similarity of these two groups in pre-treatment basic and clinical knowledge . This may mean that the outcome of the test for clinical knowledge after the problem solving treatment varies as a function of both basic and clinical knowledge before the treatment.

In fact, it appears that basic knowledge for both groups, by being very low, could not contribute to any success in the manipulation of the complex variables in the simulation program. On the contrary, it harmed the process. Maier (1970) describes the difficulties that may arise from an inadequate state of knowledge at the problem solving process. He says that the process may be blocked or impeded as selection from given alternatives is difficult, the right response components may have low priority, the final response may be to be selected without knowledge of all alternatives and finally stress may arise. Unfortunately, the present study could not reveal evidence to support theoretical considerations by Maier (ibid), Greeno (1976) and others about the role of different levels of pre-treatment basic knowledge on the effects of the problem solving approach.

The role of pre-treatment clinical knowledge upon the effects of the simulation program on the final clinical knowledge is not shown clearly either because, again, it was quite low for both classes. If this initial clinical knowledge had been better, perhaps the consequences of poor basic knowledge could have been avoided, for it is evident that well developed clinical problem solving skills are not likely to be shaken by uncertainties in basic knowledge during a single problem solving attempt.

Whatever may be the case, it seems reasonable to conclude that the simulation program failed to enhance initial learning on basic material because this was what it largely needed as a pre-requisite. It failed to enhance initial learning in clinical material because of inadequate clinical knowledge and of inadequate basic knowledge before the treatment.

In summarizing the findings of the present study, the tutorial program can be believed to guarantee a considerable success on initial subject matter learning, without requiring high levels of pre-existing knowledge. Further, it is significantly superior to the simulation in its effectiveness on both types of material.

It is possible that a later post-test, in which long term retention of the learned material might be investigated, might change the relationship between the tutorial and the simulation effects in favor of the latter. As Worthen (1968) reports, in a study of 'discovery' and 'guided' learning, although the latter method was significantly superior than the former on the test of initial learning, the retention test given after 5 weeks showed that the scores of the learning treatment group had become so low that they were surpassed by the scores of the discovery treatment group (which had not diminished considerably). However, in the case of the present study, a retention test would not reveal similar results for the simple reason that the problem solving treatment did not make performance in either type of material better than it was before exposure. Therefore, the only possible outcome at a retention test would be a diminished difference between the two treatments as the effect of the tutorial treatment would deteriorate. In no way though would the problem solving approach come to be superior than the other. It seems, then, a retention test would most probably not provide any interesting outcome.

On the Comparison Between High and Low Performance Subjects.

The present study has provided another interesting finding with respect to the high performance and low performance subjects of the fourth year group described in Chapter IV (a similar analysis could not be done for the second year subjects because of the small number of the initial four groups.). The findings suggest that the two programs used in this study do not have the same effects upon all students. This is consistent with many results in the discovery learning literature (e.g., Maier & Burke, 1966). A major finding in the present comparison was that low performance students benefit more than high performance students from the tutorial program in both types of knowledge, and are less affected by the negative effect of the simulation program. Of course, the latter program cannot be viewed as particularly harmful even on clinical material after the different tests that were used are taken into consideration. But even in this case the high performance group still remains negatively affected, since the original decline was extremely strong for clnical knowledge. The general finding is that, although the two performance groups begin with a large difference in mean scores--especially in the clinical material-- they end up with highly similar scores, and this is most obvious for the tutorial treatment. Before attempting to interpret this relationship between the two programs and the two performance groups, the fact that these two groups differed mainly with respect to clinical knowledge, deserves an elaboration. A possible reason for this is that basic knowledge on the particular subject matter had been obtained a long time ago and perhaps was refreshed at the Basic Science Option period. The recent refreshment and the nature of the material which could be remembered by mere memorization may have contributed to a lack of difference in knowledge between the two groups. On the other hand, clinical knowledge--of a general nature--was the main task of the fourth year students for a whole period of two years. This period seems to have been enough for differences in performance between good and poor students to manifest themselves. Further, an assumption that will be discussed later on is that the type of knowledge (clinical knowledge) is better possessed by the high ability students.

Now let us examine the reasons for the different effects that the tutorial program had on the two performance groups. The fact that poor students benefitted more, supports the general contention that the tutorial learning environment is more acceptable by low performance students. Thies _et al_ (1969) found a more positive attitude toward a CAI program by poor performers than by high performers, and gave the following interpretation which is also based on students' reports. High performance students as these authors assert, go further than simply giving stereotyped answers to multiple choice questions and concentrate on completing the program, 'beating' the machine' rather than attempting to learn the material presented to them. In fact, such students are believed to be antagonized by having to provide answers of this simple type. Further, at the time of the post test, according to McGuire's contentions better students often attempt to construct an answer by means of a logical reconstruction, instead of trying to recall it or recognize it (McGuire 1963). In contrast, it is poor students that usually adopt the latter way, that is, using the knowledge obtained from the teaching approach (tutorial program) by visual or auditory recall. This is less likely to have happened with basic material which, as the data has shown, was improved significantly for the high performance group too. Here though, we may assume that high performance students learned by using congnitive processes more complex than mere memorization. Interestingly enough, poor students show to have benefitted very much even from the clinical material of the program--even surpassing the initially much superior performance of the better students--thus indicating that the tutorial program may really have the capacity to teach subject matter clinical material to poor students. In this case, though, it may turn out that the low performance group achieves a very good post-treatment clinical performance by merely memorizing the material of the program, while the high performance group remains at a high performance level which is due to general clinical knowledge.

Finally, we come to the different effects of the simulation program. Although these are not as obvious as in the tutorial, the reasons that account for it seem to be the same. Basic knowledge has not undergone any major change in either group after the problem solving treatment. The small decline of the high performance group may be due to the overall confusion that the inability to solve the problem caused. This confusion may not have been occurred to the lower performance students and therefore, basic knowledge remained as it was. In fact, here too, the better students tend to question more on the interrelations of the variables in the problem solving process and get more confused--because of the lack of basic knowledge. This may also explain the drop in the clinical test. Poor students on the other hand, perhaps do not get very deeply into the complex aspects of the problem and thus do not doubt about their existing clinical skill. Perhaps, too, they do not have a lot to lose.

APPENDIX 1

A SAMPLE OF THE TUTORIAL COMPUTER-BASED PROGRAM

Two hours post-admission the lab values obtained are:

blood gases pH=7.40, $pCO2=32$, $aHCO3=18$
electrolytes $Na^{+}=138$, K=3.8, $Cl^{-}=100$, $tCO2=19$
serum salicylate 135 mg/%

Treatment with 5% gluc/saline at 250 ml/hr has resulted in a urine output of 40 ml/hr. No other treatment was given. The urine pH was 6.0. It is decided to treat with 500 mg acetazolamide IV to alkalinize the urine further.

The lab value which is the STRONGEST argument AGAINST using acetazolamide at this time is:

---

1. the serum K
2. the serum salicylate
3. the bicarbonate
4. the blood pH
5. the serum sodium

?
3

YOUR CHOICE:
3. the bicarbonate

COMMENT ON YOUR CHOICE:

The bicarbonate reflects metabolic acidosis and MAY set the sage for masked acidemia on acetazolamide administration. However, the serum bicarbonate by itself does not necessarily stay the administration of acetazolomide.

Press RETURN to continue:

Now let's review the choices:

Which is the lab value that is the strongest argument AGAINST using the acetazolamide at this time?

THE CHOICE:
1. the serum K

COMMENTARY:

The serum K is indirectly affected by the acidemia induced by acetazolamide administration. Serum potassium would rise and urinary potassium clearance would increase. This would eventually result in potassium depletion. This would not be the major argument against using acetazolamide at this time.

Press RETURN to continue:

APPENDIX 2

A SAMPLE OF THE SIMULATION COMPUTER-BASED PROGRAM

CASE PRESENTATION

A 56 year old stockbroker is admitted to the intensive care unit, where you are the physician on duty.His secretary tells you that the patient has been 'bearish' on the stock market lately.

One hour before admission,the patient said "that he was going to end this headache once and for all".Shortly thereafter,he was found at his desk in a dazed state.

He was nauseated, but had not vomited. His secretary noted an empty bottle of aspirin (100 tablets),which shehad bought for him that morning.

In emergency,lavage had been performed with return of about 15 mls. of chalky white material identified as acetylsalicylic acid.

Press RETURN to continue:

*** TIME POST-ADMISSION IS 0 HOURS, 0 MINUTES ***

Electrolytes	Blood Gases	ASA Level

The patient is drowsy but responsive

Please enter your next request: _

*** TIME POST-ADMISSION IS 0 HOURS, 0 MINUTES ***

Electrolytes	Blood Gases	ASA Level

Please enter your next request: co_

*** TIME POST-ADMISSION IS 0 HOURS, 0 MINUTES ***

Electrolytes	Blood Gases	ASA Level

Choice List

Heart rate/pulses (hr)	CVP (cv)	Glucose in water (gw)
Blood Pressure (bp)	Urine Output (uo)	Glucose in saline (gs)
Resp. rate/vol. (rv)	Urine pH (up)	Normal saline (ns)
Cardiac monitor (mo)	Input/Output (io)	2/3 - 1/3 (23)
Depth of coma (co)	Defibrillate (df)	Isotonic bicarb (ib)
Hypertonic bicarb (hb)	KCl in I.V. (K+)	Restore Screen (rs)
Electrolytes (ly)	Acetazolamide (az)	Go on (go)
Blood Gases (bg)	Review (re)	Quit (qu)
ASA level (as)	Print Record (pr)	

Press RETURN to continue: _

*** TIME POST-ADMISSION IS 2 HOURS, 5 MINUTES ***

Electrolytes (00:00)	Blood Gases (00:00)	ASA Level (00:00)
Na+ = 140.0 mEq/l	pH = 7.5	sASA = 121.7 mg%
K+ = 3.5 mEq/l	HCO3- = 21.7 mmol/l	
Cl- = 98.6 mEq/l	pCO2 = 30.8 mm Hg	
CO2 = 22.7 mmol/l		

Lab results are now back....

_

Please enter your next request: _

*** TIME POST-ADMISSION IS 6 HOURS, 5 MINUTES ***

Electrolytes (00:00)	Blood Gases (00:00)	ASA Level (00:00)
Na+ = 140.0 mEq/l	pH = 7.5	sASA = 121.7 mg%
K+ = 3.5 mEq/l	HCO3- = 21.7 mmol/l	
Cl- = 98.6 mEq/l	pCO2 = 30.8 mm Hg	
CO2 = 22.7 mmol/l		

***The patient is apneic and has no pulse!

How many minutes to your next decision point? 120

Press RETURN to continue:

APPENDIX 3

THE LIST OF POSSIBLE REQUESTS AVAILABLE AT THE SIMULATION COMPUTER-BASED PROGRAM.

PHYSICAL EXAM AND NURSING FINDINGS

Heart rate and pulses (hr)

Blood pressure (bp)

Respiratory rate/valume (rv)

Cardiac monitor (mo)

Depth of coma (co)

Central venous pressure (cv)

Hourly urine output (uo)

Urine pH (up)

Cumulative input/output (io)

TREATMENTS

Glucose in water (gw)

Glucose in normal saline (gs)

Normal saline (ns)

3-1/3% glucose in 0.3n saline (23)

Isotonic NaHC03 in water (ib)

Hypertonic NaHC03 in water (hb)

Acetazolamide (DIAMOX) (az)

KCl added to I.V. (k+)

Defibrillation (df)

LAB TESTS

Electrolytes (ly)

Blood gases (bg)

Serum asa level (as)

Print record (pr)

SPECIAL FUNCTIONS

Treatment review (re)

Restore screen (rs)

Go on to next decision point (go)

Quit (qu)

APPENDIX 4

TESTS USED IN THE EXPERIMENT: TEST A

STUDY ON EFFICIENT APPROACHES

FOR TEACHING CLINICIANS

Please check (✓) the answer that, to your opinion, best fits each statement below:

Please select choice (A) if (A) is greater than (B)
Please select choice (B) if (B) is greater than (A)

1. A. The urinary clearance of salicylate at a urine pH of 5.5 and urine output of 100 ml/hr.
   B. The urinary clearance of salicylate at a urine pH of 7.5 and urine output of 10 ml/hr.

2. A. The risk of cardiac arrhythmias at a blood pH of 7.30 and K^+_s=4meq/L
   B. The risk of cardiac arrhythmias at a blood pH of 7.5 and K^+_s=4meq/L

3. A. The minute volume respir. rate x tidal volume) in a human with a blood pH of 7.30 and a serum salicylate of 10 mgm%
   B. The minute volume (resp. rate x tidal volume) in the same human (same age) with a blood pH of 7.30 and a serum salicylate of 45 mgm%

4. The administration of acetazolamide carries which of the following risks?

   1. Rapid urinary loss of K^+ and hypokalemia
   2. Inhibition of hepatic carbonic anhydrase
   3. Systemic acedemia with increased CNS toxicity
   4. Increased urinary pH with salicylate crystalluria

Please select a letter:

A=1,2,3; B=1,3; C=2,4; D=4; E=none of the above

5. Na^+=135meq/L, K^+=4.0meq/L, Cl^-=96meq/L, tCO_2=16meq/L
   Pick the acid-base disorder (s) that might result in the electrolytes given above:

   1. metabolic alkalosis
   2. respiratory alkalosis
   3. respiratory acidosis
   4. metabolic acidosis

Please select a letter:

A=1,2,3; B=1,3; C=2,4; D=4; E=none of the above

6. In monitoring bicarbonate therapy of salicylate intoxication several principles apply. Pick it (them).

   1. urinary pH must be performed every ½hour
   2. the formulas for calculating bicarbonate replacement are only guidelines
   3. blood gases must be determined at least every hour
   4. serum salicylate determinations every hour are necessary

Please select a letter:

A=1,2,3; B=1,3; C=2,4; D=4; E=none of the above

7. A patient with salicylate intoxication with a serum potassium of 7.0 meq/L and blood pH of 7.20 would best be treated with:

1. Na^+K^+ exchange resin (extracts body K^+) and bicarbonate
2. acetazolamide
3. hemodialysis
4. 5% glucose at 500 ml/hr

Please select a letter:

A=1,2,3; B=1,3; C=2,4; D=4; E=none of the above

---

8. The metabolic acidemia is often seen in untreated salicylate intoxication
BECAUSE
salicylate ion is an acidic ion

Please choose: A=TT rel
B=TT no rel
C=TF
D=FT
E=FF

9. A high urine output increases the renal clearance of salicylate
BECAUSE
salicylate ion is not very soluble in urine

Please choose: A=TT rel
B=TT no rel
C=TF
D=FT
E=FF

---

10. One ampoule of bicarbonate contains: (please choose one)

A. 44meq of HCO_3 and 50 ml water
B. 44meq of Na^+,44meq of HCO_3, and 50 ml of water
C. 44meq of Na^+,44meq of lactate, and 50 ml of water
D. 50meq of Na^+,50meq of HCO_3, and 50 ml of water
E. none of the above

11. At time A (2 hours post ingestion) in a case of salicylate intoxication the biochemistry is:

Na^+_s-135 meq/L	pH-7.40
K^+_s -3.8 meq/L	pCO_2-35mm Hg
Cl^--90 meq/L	HCO_3-18mmol/L
$tHCO_3$-20 meq/L	sASA-95mgm%

At time A, treatment with acetazolamide 500 mg IV* and 5% g/w at 250ml/hr is begun.

2 hours after time A the patient is suddenly considerably more stuporous with apnoic spells and only slight response to deep pain. After drawing blood for biochemistry again, pick from the following list the 2 THINGS which you would do next IN ORDER OF EXECUTION. (assume above Rx continues)

* IV= Intravenously

A. stop 5% g/w (glucose in water)
B. give 4 ampoules of bicarbonate
C. check the cardiac monitor
D. call the renal resident to arrange hemodialysis
E. prepare for tracheal intubation and assisted ventilation
F. something else (specify)

12. 8 hours after appropriate treatment, the situation described in Q.11 has improved considerably. Respiration is full and spontaneous and the patient's main complaint is marked deafness. The latest biochemistry reads:

NA+ -142 meq/L	pH-7.43
K+-2.8 meq/L	PCO_2-34mm Hg
Cl^--98 meq/L	$aHCO_3$-16 meq/L
tCO_2-18 meq/L	sASA-44mgm%

The patient suddenly stiffens, arches his back and turns dusky. The cardiac monitor is indecipherable because of movement. The first 3 THINGS you would do IN ORDER OF EXECUTION are:

A. inject diazepam - 10 mg intravenously
B. give K^+ - 40 meq p.o.
C. give K^+ - 40 meq / 30 min IV
D. defibrillate + assist ventilation with endotrachea intubation
E. give 2 ampoules (88meq) of HCO_3
F. defibrillate + assist ventilation with an Ambu bag and face mask
G. inject lidocaine - 75 mg IV @ 10 minutes
H. give K^+ - 80meq /30 min. IV
I. give morphine, furosemide O_2 and apply rotating tourniquets
J. something else (specify)

13. 12 hours after a 75-tablet aspirin ingestion the biochemistry reads:

Na^+-148 meq/L	pH-7.42
K^+-3.7 meq/L	pCO_2-38mm Hg
Cl^--102 meq/L	$aHCO_3$-19mg/L
$tHCO_3$-20 meq/L	sASA28mg%

In the last 2 hours the urine output formerly at 150-200 ml/hr has fallen to 40 ml/hr in spite of at most 500 ml/hr of various intravenous fluids. Respiratory rate has increased to 35/min and sounds laboured. The patient appears dusky. He develops a persistent cough and suddenly the same thing happens as in the previous question.

Please answer as in the previous question from the choices given there.

---

14. Please, select choice (A) if (A) is greater than (B)
select choice (B) if (B) is greater than (A)
select (C) if (A)=(B)

A. The absolute change in blood pH that would lower the serum potassium from 4.0 to 3.0 meq/l.
B. The absolute change in blood pH that would raise the serum potassium from 4.0 to 5.0 meq/l.

---

15. Stomach lavage 6 hours after ingestion of 80 tablets of ASA is very unlikely to remove significant amounts of acetylsalicylic acid

BECAUSE

regular doses of acetylsalicylate are rapidly absorbed in the stomach 'speeding' to the 'pain centres' via the circulation.

Please choose: A=TT rel
B=TT no rel
C=TF
D=FT
E=FF

APPENDIX 5

TESTS USED IN THE EXPERIMENT: TEST B

STUDY ON EFFICIENT APPROACHES

FOR TEACHING CLINICIANS

Please choose the answer that, to your opinion, best fits each statement below:

1. 1000 ml of "2/3 - 1/3" solution contains:

   1. 33.3 gm of glucose
   2. 66.7 gm of glucose
   3. 0.30 meq of Na^+
   4. 0.33 meq of Na^+

Please select a letter:

A=1,2,3; B=1,3; C=2,4; D=4; E=none of the above

2. The increased minute volume seen in patients with established (> 4 hrs post-ingestion) salicylate intoxication is caused by:

   1. Central(CNS) respiratory stimulation by salicylate
   2. increased ventilation seen reflexly in nausea
   3. chemoreceptor (CNS) respiratory stimulation by metabolic acid
   4. salicylate induced depression of sensitivity of pulmonary stretch-receptors

Please select a letter:

A=1,2,3; B=1,3; C=2,4; D=4; E=none of the above

---

3. Lactic acid is the main metabolic acid produced in the later stages of salicylate intoxication

BECAUSE

lactic acid is the main metabolite of salicylic acid.

Please choose:

A=TT rel
B=TT no rel
C=TF
D=FT
E=FF

4. Hemodialysis (against normal dialysate) is the most rapid way known of eliminating salicylate

BECAUSE

salicylate is more ionized in the dialysate (which is alkaline relative to blood)

Please choose:

A=TT rel
B=TT no rel
C=TF
D=FT
E=FF

5. Before starting acetazolamide treatment in ASA intoxication, it is important to ascertain that the blood pH is greater than 7.4

BECAUSE

acetazolamide's effect on carbonic anhydrase is opposed by acidemia.

Please choose:
A=TT rel
B=TT no rel
C=TF
D=FT
E=FF

Please select choice (A) if (A) is greater than (B)
select choice (B) if (B) is greater than (A)
slect (C) if (A)=(B)

6. A. The rate of urinary HCO_3^- elimination 1 hour after a dose of 250 mg of acetazolamide IV*.

   B. The rate of urinary HCO_3^- elimination in the same patient 1 hour after 500 mg acetazolamide IV.

7. A. The size of the increase in renal salicylate clearance seen, when the urine pH goes from 6.5 to 7.0.

   B. The size of the increase in renal salicylate clearance (in the same situation) when the urine pH goes from 7.0 to 7.5.

---

8. A 42-year-old depressed patient has been found unconscious at home. He appears to have carried out his oft repeated threat to TV commercials "take no more of their @!#!". Instead he has taken a non-advertised brand of aspirin, thereby fully disagreeing (to the tune of 75 adult tablets) with the current philosophy that "all aspirin is not really alike". The time of ingestion is not clear, but can be narrowed down to the last 48 hours.
Although he has been seeing a psychiatrist, he has been on no medications according to his family. Gastric lavage yields 1 tbsp of white sludge identified as ASA. The patient responds (by withdrawing) to deep pain. His vital signs are stable but ventilation is deep and sighing. After getting this history and these physical signs, pick IN ORDER OF EXECUTION the <u>4 things</u> you would do next:

   A. call his psychiatrist to see if he has taken ASA overdose before
   B. call his psychiatrist to make sure patient wasn't on a sedative or antipsychotic agent
   C. draw blood for blood gases, electrolytes
   D. draw blood for serum salicylate
   E. start "2/3-1/3" @ 500 ml/hr
   F. give 4 ampules of bicarbonate
   G. give 500 mg of acetazolamide intravenously
   H. start isotonic $NaHCO_3^-$ solution at 500 ml/hr
   I. start K^+ at 40 meq/hr intravenously
   J. give K^+ - 50 meq over 10 minutes
   K. give 2 ampules of hypertonic bicarbonate
   L. something else (specify)

* IV = Intravenously

9. In your treatment of the patient introduced in question #8, you give 500 ml/hr of saline over the first 4 hours, because you read somewhere that salicylate inoxications need fluid replacement and that sodium may prevent the appearance of acidemia. However, total urine output is 50 ml over 4 hours. The blood pressure has been hovering at 100/70 and the patient is considerably more confused.
Pick IN ORDER OF EXECUTION the 3 steps you would now take from the list in Question # 8.

10. 1/2 hour after the situation described in question # 9 you get electrolytes and blood gases which read:

Na^+-152 meq/l	pH-7.22
K^+-4.0 meq/l	pCO_2-28 mm Hg
Cl^--90 meq/l	$aHCO_3$-11 mmol/l
tCO_2-12 meq/l	

While you are pondering these, the patient begins to gasp and turn blue. The cardiac monitor shows ventricular fibrillation. After you successfully defibrillate him and start assisted ventilation, your next <u>three</u> steps IN ORDER OF EXECUTION would be: (pick from these:)

A. give 4 ampoules of $NaHCO_3$
B. give furosemide - 10mg intravenously
C. get a serum glucose and administer hypertonic glucose
D. start K^+ @ 40 meq/hr intravenously
E. give acetazolamide - 500 mg intravenously
F. give digoxin - 0.25mg IV
G. order serum salicylate
H. give 2 ampules of $NaHCO_3$
I. get blood for gases and electrolytes
J. something else (specify)

---

11. Pick the factor(s) directly governing renal clearance of salicylate:

1. urine potassium
2. serum pH
3. serum chloride
4. urine output

Please select a letter:

A=1,2,3; B=1,3; C=2,4; D=4; E=all of the above.

12. The risks of bicarbonate therapy in salicylate intoxication include:

1. sodium and water overload
2. alkalemia depressing respiration (effect opposite to that of acidemia)
3. hypokalemia and cardiac arrhythmias
4. conversion to CO_2 with CO_2 narcosis

Please choose a letter:

A=1,2,3; B=1,3; C=2,4; D=4; E=all of the above

Please, select choice (A) if (A) is greater than (B)
select choice (B) if (B) is greater than (A)
select (C) if (A)=(B)

13. A. The rate of renal potassium loss in a patient treated with acetazolomide (500mg) and blood pH=7.35

B. The rate of renal potassium loss in the same patient on the same treatment with a blood pH=7.45

14. A. The per cent increase in renal clearance of salicylate when the urine pH is increased from 5.5 to 7.5.

B. The per cent increase in renal clearance of salicylate metabolites (glucuronides, homogentisic acid conjugates) when the urine pH is increased from 5.5 to 7.5 in the same patient.

---

15. The clinical severity of a salicylate intoxication can be gauged best by: (please choose one)

A. the degree of salicylism (tinnitus + deafness)
B. 2 separate salicylate determinations 2 hours apart and a Done nomogram
C. an estimate of duration of exposure to high salicylate and high blood hydrogen ion concentrations
D. combination of A and B
E. combination of B and consultation with a clinical pharmacologist

APPENDIX 6

QUESTIONNAIRE ASSIGNED BEFORE EXPERIMENTAL TREATMENT

<u>PART II</u>

Name:________________________________

Please answer:

1. Have you ever used the 'ASA' simulation CAI program before?

YES ☐ NO ☐

If YES, when was the last time?

one year ago or less ☐ more than one year ago ☐

2. Between our first meeting and this one, did you happen to study, or attend any lecture, on the material that is involved in this study?

YES ☐ NO ☐

APPENDIX 7

LETTER SENT TO SECOND-YEAR STUDENTS REQUESTING THEIR PARTICIPATION.

4th April 1978

Dear Student:

The Centre for Medical Education in co-operation with the Department of Pharmacology is currently involved in an experiment, evaluating the computer aided instructional program called 'ASA' (aspirin intoxication treatment).

In order for this project to proceed we require the assistance of a number of second year students and it is hoped that through the contact of this letter you will be willing to participate in our project.

The amount of time required over the two sessions of your involvement, would be 15 minutes for the first and approximately one hour for the second session. Besides, the content of the computer program might be an opportunity for you to improve your knowledge of acid-base physiology.

It will be appreciated if you call by phone to either of the undersigned at your earliest convenience. Tel.: 392-4928.

Thanking you.

Sincerely

Evie Tsouna
Project Coordinator

Hugh M. Scott, M.D.
Director

APPENDIX 8

% MEAN SCORES OF FOURTH YEAR STUDENTS FAMILIAR WITH THE SIMULATION PROGRAM. (MEAN SCORES FOR THE WHOLE GROUP ARE IN PARENTHESES)

TREATMENT	PERFORMANCE IN	BEFORE		AFTER	
	Overall	51.3	(50.0)	65.0	(62.5)
TUTORIAL	Basic	52.5	(45.3)	68.9	(70.4)
	Clinical	50.3	(52.3)	63.3	(59.0)
	Overall	60.6	(53.2)	45.6	(45.8)
SIMULATION	Basic	46.0	(45.8)	47.3	(46.0)
	Clinical	67.4	(56.7)	45.0	(45.6)

APPENDIX 9

ANALYSIS OF VARIANCE TABLES FOR OVERALL, 'BASIC' AND 'CLINICAL' PERFORMANCE OF FOURTH YEAR STUDENTS.

ANALYSIS OF VARIANCE ON THE OVERALL PERFORMANCE FOR THE TWO TESTS AND THE TWO TREATMENTS.

SOURCE OF VARIATION	SS	DF	MS	F	SIGNIFICANCE
Main Effects	1374.7	2	687.3	7.29	.001
Treatment	1213.4	1	1213.4	12.87	.001
Before-after	161.3	1	161.3	1.71	.191
Interaction					
Treat x Before-after	2841.8	1	2841.8	30.13	.001
Error	9807.4	104	94.3		

ANALYSIS OF VARIANCE ON PERFORMANCE IN BASIC QUESTIONS FOR THE TWO TESTS AND THE TWO TREATMENTS.

SOURCE OF VARIATION	SS	DF	MS	F	SIGNIFICANCE
Main Effects	8115.4	2	4057.7	15.49	.001
Treatment	3826.0	1	3826.0	14.60	.001
Before-after	4289.3	1	4289.3	16.37	.001
Interaction					
Treat x before-after	4202.7	1	4202.7	16.04	.001
Error	27243.2	104	261.9		

ANALYSIS OF VARIANCE ON PERFORMANCE IN CLINICAL QUESTIONS

FOR THE TWO TESTS AND THE TWO TREATMENTS.

SOURCE OF VARIATION	SS	DF	MS	F	SIGNIFICANCE
Main Effects	672.6	2	336.3	2.13	.122
Treatment	549.1	1	549.1	3.47	.062
Before-after	123.5	1	123.5	.78	.999
Interaction					
Treat x Before-after	2137.4	1	2137.4	13.50	.001
Error	16454.7	104	158.2		

APPENDIX 10

ANALYSIS OF VARIANCE TABLES FOR OVERALL, 'BASIC' AND 'CLINICAL' PERFORMANCE OF SECOND YEAR STUDENTS.

ANALYSIS OF VARIANCE ON THE OVERALL PERFORMANCE FOR THE TWO TESTS AND THE TWO TREATMENTS.

SOURCE OF VARIATION	SS	DF	MS	F	SIGNIFICANCE
Main Effects	1419.6	2	709.8	5.82	.005
Treatment	511.8	1	511.8	4.19	.041
Before-after	907.8	1	907.8	7.44	.008
Interaction					
Treat.x Before-after	1695.9	1	1695.9	13.89	.001
Error	10739.4	88	122.0		

ANALYSIS OF VARIANCE ON PERFORMANCE IN BASIC QUESTIONS FOR THE TWO TESTS AND THE TWO TREATMENTS.

SOURCE OF VARIATION	SS	DF	MS	F	SIGNIFICANCE
Main Effects	1888.1	2	944.1	3.25	.042
Treatment	829.2	1	829.2	2.85	.091
Before-after	1058.9	1	1058.9	3.65	.056
Interaction					
Treat.x Before-after	1024.3	1	1024.3	3.53	.060
Error	25527.2	88	290.0		

ANALYSIS OF VARIANCE ON PERFORMANCE IN CLINICAL QUESTIONS FOR THE TWO TESTS AND THE TWO TREATMENTS.

SOURCE OF VARIATION	SS	DF	MS	F	SIGNIFICANCE
Main Effects	1221.7	2	610.8	3.38	.037
Treatment	391.7	1	391.7	2.17	.140
Before-after	830.0	1	830.0	4.59	.033
Interaction					
Treat x Before-after	2067.4	1	2067.4	11.45	.001
Error	15887.5	88	180.5		

BIBLIOGRAPHY

Barnoon, S. and Wolfe, H. *Measuring the Effectiveness of Medical Decisions*, Charles C. Thomas, Publisher, 1972.

Barrows, H.S. and Abrahamson, S. The Programmed Patient: A Technique for Appraising Student Performance in Clinical Neurology. *Journal of Medical Education*, August 1964, *39*, 802-805.

Barrows, H.S., Feighner, J., Newfeld, V., and Norman, G. *An Analysis of the Clinical Methods of Medical Students and and Physicians*. Report submitted to the Province of Ontario Department of Health and Physicians Services In., 1978.

Barrows, H.S., and Mitchell, D.L.M. The 'Problem-Box' Approach to Learning and its Effect on the Undergraduate Curriculum. *Eleventh Annual Conference on Research in Medical Education*, Nov. 1972, 160-165.

Berner, E.S., Hamilton, L.A., and Best, W.R. A New Approach to Evaluating Problem-Solving in Medical Students. *Journal of Medical Education*, July 1974, *49*, 666-672.

Bitzer, D.L. PLATO. In L.M. Stolurow, T.I. Peterson and A.C. Cunningham (eds), *Computer Assisted Instruction in the Health Professions*, Enttelek, Newburyport, Massachusetts, 1970, 177-194.

Brown, J.S. and Burton, R. Multiple Representation of Knowledge for Tutorial Reasoning. In D. Boborow & Collins (eds.) *Representation and Understanding*. New York, Academic Press, 1975.

Craig, R.C. Directed vs Independent Discovery of Estabished Relationships, *Journal of Educational Psychology*, 1956, *47*, 223-224.

De Groot, A.D. *Thought and Choice in Chess*. New York: Basic Books, 1965

Elstein, A., Kagan, N., Shulman, L., Jason, H. and Loupe, M. Methods and Theory in the Study of Medical Inquiry. *Journal of Medical Education*, Feb 1972, *47*, 85-92.

Elstein, A.S., Shulman, L.S. and Sprafka, S.A. *Medical Problem Solving: An Analysis of Clinical Reasoning*. Cambridge, Mass., Harvard University Press, 1978

Feurzeig, W. Decision Maing. In L.M. Stolurow, T.I. Peterson and A.C. Cunningham (eds.), *Computer Assisted Instruction in the Health Professions*, Ettelek, Newburyport, Massachusetts, 1970,114-123.

Feurzeig, W., Munter, P., Swets, J., and Breen, M. Computer-Aided Teaching in Medical Diagnosis. _Journal of Medical Education_, Aug. 1964 _39_, 746-754.

Gagne, R.M., and Brown, L.T. Some Factors in the Programming of Conceptual Learning, _Journal of Experimental Psychology_, 1961, _62_, 313-321.

Goldman, R.D., and Hudson, D.J. A Multivariate Analysis of Academic Abilities and Strategies for Successful and Unsuccessful College Students in Different Major Fields. _Journal of Educational Psychology_, 1973, _65_, 364-371.

Goran, M.J., Williamson, J.W., and Gonnella, J.S. The Validity of Patient Management Problems. _Journal of Medical Education_, Feb. 1 973, _48_, 171-177.

Greeno, J.G. Cognitive Objectives of Instruction: Theory of Knowledge for Solving Problems and Answering Questions. In Klahr, E. (Ed.), _Cognition and Instruction_, New Jersey: Laurence Erlbaum Associates, 1976, 123-159.

Guthrie, J.T. Expository Instruction Versus a Discovery Method. _Journal of Educational Psychology_, 1967, _58_, 45-49.

Hammidi, I.B., and Fonkalsrud, E.W. Case Studies in Surgery. In L.M. Stolurow, T.I. Peterson, and A.C. Cunningham (eds.), _Computer Assisted Instruction in the Health Professions_, Enttelek, Newburyport, Massachusetts, 1970, 125-139.

Helfer, R.E. and Slater, C.H. Measuring the Process of solving Clinical Diagnostic Problems. _British Journal of Medical Education_, 1971, 548-552.

Hendrix, G. A New Clue to Transfer of Training. _Elementary School Journal_, 1947, _48_, 197-208.

Hoffman, P.J. The Paramorphic Representation of Clinical Judgement. _Psychological Bulletin_, 1960, _57_ (2), 116-131.

Johnson, P.E., Moller, J.H., and Bass, G.M. Analysis of Expert Diagnosis of a Computer Simulation of Congenital Heart Disease. _Journal of Medical Education_, May 1975, _50_, 466-469.

Kittel, J.E. An Experimetal Study of the Effect of External Direction During Learning of Transfer and Retention of Principles. _Journal of Educational Psychology_, 1957, _48_, 391-405.

Lefever, R.D. and Johnson, J.K. _Survey on the Use of Computers in Instruction_, Association of American Medical Colleges, Washington, D.C., 1976.

Maatsch, J.L., An Introduction to Patient Games: Some Fundmentals of Clinical Instruction. Office of Medical Education Research and Development, Michigan State University, 1974.

Maier, N.R.F., Problem Solving and Creativity in Individuals and Groups. Belmont, California: Brooks/Cole, 1970.

Maier, N.R.F. and Burke, R.J. Test of the Concept of 'Availability of Function' in Problem Solving. Psychological Reports, 1966. 19, 119-125.

Malin, J.E. An Analysis of Strategies from Solving Certain Substitution Problems. Dissertation Abstract International, Feb. 1974, 34, (8-13), 4089.

McGuire, C.H. A Process Approach to the Construction and Analyis of Medical Examinations. Journal of Medical Education, 1963, 38, 556-563

McGuire, C.H. and Babbott, D. Simulation Technique in the Measurement of Problem-Solving Skills. Journal of Educational Measurement, Spring 1967, 4, (1), 1-10.

McGuire, C.H., Solomon, L.M., and Forman, P.M. Clinical Simulations. Selected Problems in Patient Management. New York: Appleton-Century-Crofts, 1976.

Meyer, F., and Beaton, G., An Evaluation of Computer Assisted Teaching in Physiology. Journal of Medical Education, March 1974, 49, 295-297.

Miller G.A., Galanter. E., Pribram, K. Plans and the Structure of Behavior. New York, Holt. 1960

Newell, A. and Simon, H.A. Human Problem Solving. Englewood Cliffs, N.J., Prentice Hall, 1972.

O'Neill, T., Sewall, J., and Marchand, R. Time-Shared Computer-Assisted Preclinical Instruction: A Short Trial and Evaluation. Journal of Medical Education, Sept. 1976, 51, 765-767.

Palva, J. P., Measuring Clinical Problem Solving. Journal of Medical Education, 1974, 8, 52-56.

Palva, J.P. and Korhonen, V., Validity and Use for Written Simulation Tests of Clinical Performance. Journal of Medical Education, Aug. 1976, 51, 657-661.

Ray, W.S., Pupil Discovery vs Direct Instruction. Journal of Experimental Education, 1961, 29, 271-280.

Sajid, A., Lipson, L.F., and Telder, T.V. A Simulation Laboratory for Medical Education, Journal of Medical Education, Oct. 1975, 50, 970-975.

Shulman, L.S. Psychological Controversies in the Teaching of

Science and Mathematics. In H.F. Clarizio, R.C. Craig, and W.A. Mehrens, _Contemporary Issues in Educational Psychology_, Allyn & Bacon, 1974.

Shulman, L., Loupe, M., and Piper, R. _Studies of the Inquiry Process_. Educational Publication Services College for Education, Michigan State University, 1968.

Simon, H.A. Artificial Intelligence Systems that Understand. _5th International Joint Conference on Artificial Intelligence, 1977_. MIT Cambridge-Massachusetts, _2_, 1059-1073.

Sorlie, W., and Jones, L. Increasing Medical Student Performance with an Interactive, Computer-Assisted Appraisal System. _Journal of Computer-Based Instruction_, Feb. 1976, _2_ (3), 57-62.

Stolurow, L.M., Peterson, T.I., and Cunningham, A.C. (Eds.), _Computer-Assisted Instruction in the Health Professions_, Enttelek, Newburyport, Massachusetts, 1970.

Thies, R., Harless, W., Lucas, N., and Jacobson, E. An Experiment Comparing Computer-Assisted Instruction with Lecture Presentation in Physiology. _Journal of Medical Education_, Dec. 1969, _44_, 1156-1160.

Votaw, R.G. and Farquhar, B.B. Current Trends in Computer-Based Education in Medicine. _Educational Technology_, April 1978, 54-56.

Williamson, J.W., and McGuire, C. Consecutive Case Conference: An Educational Evaluation. _Journal of Medical Education_, Oct. 1968, _43_, 1068-1074.

Wittrock, M.C. Verbal Stimuli in Concept Formation: Learning by Discovery. _Journal of Educational Psychology_, 1963, _54_, 183-196.

Worthen, B.R. Discovery and Expository Task Presentation in Elementary Mathematics, _Journal of Educational Psychology Monograph Supplement_, 1968, _59_ (1), Part 2.