DOCUMENT RESUME

ED 073 613 EM 010 715

AUTHOR	Lasater, Mary Beth
TITLE	Design and Evaluation of a CAI Program in Chemical Equilibrium. Technical Report Number 8.
INSTITUTION	Texas Univ., Austin. Computer-Assisted Instruction Lab.
SPONS AGENCY	National Science Foundation, Washington, D.C.
PUB DATE	Jun 71
NOTE	45p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Chemical Equilibrium; *Computer Assisted Instruction; Computer Programs; Educational Technology; Experimental Programs; *Instructional Design; *Program Design; *Program Development; Program Effectiveness; *Program Evaluation; Systems Approach; Undergraduate Study
IDENTIFIERS	*Coursewriter III

ABSTRACT

The design, development, and evaluation of a computer-assisted instruction (CAI) program in first-year college chemistry is described in this report. The program was centered around the concept of chemical equilibrium and used a systems approach with explicitly stated guidelines. The purpose of this experiment was not to measure the effectiveness of the course content, but rather to investigate the usefulness and effectiveness of the particular developmental processes and design techniques that were employed. The results of the program revealed basically that students with low mathematical capabilities upon entering the course take longer to finish the prescribed sequences and that a student's criterion score is inversely related to his path length through the course. (MC)

THE UNIVERSITY OF TEXAS AT AUSTIN

Computer Assisted Instruction Laboratory

AUSTIN

DESIGN AND EVALUATION OF A CAI PROGRAM IN CHEMICAL EQUILIBRIUM

TECHNICAL REPORT NO. 8

Mary Beth Lasater

June 1971

Supported By:

THE NATIONAL SCIENCE FOUNDATION
Grant GJ 509 X

Computer-Assisted Instruction Laboratory
The University of Texas at Austin
Austin, Texas 78712

INTRODUCTION

The use of computer-assisted instruction (CAI) as an innovative medium for individualized instruction and as a form of instrumentation for educational research is widely recognized. There have been, however, few efforts to develop materials that have a course-related content and also have explicit instructional guidelines and easily manipulated variables. This seems unfortunate, since findings obtained in the learning laboratory under experimental conditions and with contrived or unrealistic content may not be generalizable to the actual classroom or to an individualized instructional situation. In the development of this program an attempt has been made to include both of these potential contributions of computer-assisted instruction. The program itself is justified primarily in terms of the instructional needs of specific students; at the same time, both the content itself and various of the instructional features were influenced by considerations of research utility.

The Goal of Science Education

The study of science should develop the student's ability to apply scientific generalizations to a variety of different situations. Even for those students whose primary interest lies in the liberal arts, such applications in the physical sciences at the college level should be, in part at least, quantitative in order to reflect adequately this aspect of science. For the student who is majoring in one of the sciences or in engineering, these quantitative applications assume a much greater significance in the introductory college courses because both the remainder of his preprofessional study and also his professional activities will require use of them.

Within the field of chemistry, there are numerous principles and and generalizations that form the bases for quantitative applications. One group of these concerns *chemical equilibrium*, a topic which is generally included in first-year college chemistry and especially in those courses designed for students who will pursue further study in science. At The University of Texas at Austin, *dynamic equilibrium* is one of the areas accorded major emphasis in freshman chemistry, and approximately five to six weeks of lecture time are usually allowed for it.

Two factors in addition to its relevance for students of first-year college chemistry were considered in choosing the content for this particular program. In the first place, the application of equilibrium principles appears to be typical of a variety of quantitative applications. Also, the terminal objectives are clear-cut and readily definable, and the task analysis is a relatively straightforward one.

Instructional Needs of Science Students

If students are to acquire skill in making quantitative applications of scientific principles, it is desirable for them to practice such applications for themselves. On the other hand, guidance and additional instruction should be available to the student when they are needed. In the past such supervised practice has frequently been provided through small-group discussion sessions separate from the regular lectures. Many institutions, including The University of Texas at Austin, have now discontinued this practice due to the scheduling difficulties encountered with large numbers of students and to the shortage of suitable instructional space and of qualified personnel. Instruction in quantitative applications and in the handling of numerical problems has, therefore, been incorporated into the lecture sections, each of which ordinarily enrolls two hundred or more students. Thus, too much of the responsibility for actual practice in these skills must be placed upon the students themselves.

First-year students in college chemistry, even those who are planning to major in one of the sciences or in engineering, typically present a rather wide variety of backgrounds in science and mathematics as well as a range of individual aptitudes. These differences become particularly apparent in numerical applications where not only chemical concepts but also mathematical skills must be called upon. This suggests the need for *individualization*, where this term is used somewhat broadly to describe any type of curricular adjustment that is based upon learner characteristics.

Computer-Assisted Instruction

One possible way of meeting the student's need for individualized instruction is through the use of CAI. Such instruction permits the student to advance at his own pace until a criterion level of performance is reached and to do so, to a large extent, at the time of his choice. It also offers a number of other advantages for the student, advantages that are generally unavailable with alternative modes of instruction. In the first place, the student proceeds through the program along a pathway that is directly related to his own responses. This branching can be made very complex and extensive with computer-assisted instruction because the computer can be programmed to make decisions based not only upon a particular response to a specific item but also upon cumulative responses or response patterns to a series of questions or problems. In addition, the assistance given to the student can often be made quite explicit, since the author who is also experienced as an instructor can anticipate many of the errors the student is likely to make and can provide guidance that is directly related to a particular incorrect response. This avoids, for the student, the frustration of knowing that an answer is wrong but having no real clue about the source of error. A further instructional advantage results because the student can be encouraged, even forced, to proceed on his own as far as he is able; thus, he is more likely to develop real ability to apply principles rather than merely skill in recalling a series of examples. The same flexibility that produces multiple pathways through the program has been found to be highly motivating for many students. They feel that the instructional materials, rather than being impersonal and mechanical, are responding to their individual needs and difficulties. A final, potentially important, advantage is associated with the use of the computer in science instruction. Because these devices have widespread applications in the sciences and in engineering, early experience with them is valuable to the student who is planning to enter one of these fields and who will, therefore, probably be called upon to make more direct use of them later in his study.

In addition to its advantages for the student, a computer-assisted instructional program of this type can also have potential usefulness in educational research. If this potential is to be realized, the design process used to produce the instructional materials should be clearly specified and the program itself should have identifiable instructional variables that can be manipulated at a later time. An attempt has been made to incorporate these features into the computer-assisted program that was developed for this study.

Purpose of the Study

Because of the instructional needs of students and because of the possible utility of computer-assisted instruction in meeting these needs, a program designed to help students in first-year college chemistry acquire skill in applying the principles of chemical equilibrium to problems involving aqueous solutions has been developed and evaluated. An instructional systems approach was used in developing the program, and explicitly stated guidelines were followed in the preparation of the materials themselves. The final program was written in Coursewriter III for use on a time-sharing basis with an IBM 360/50 computer system. In addition to the developmental testing with individual learners and very small groups, the final computer-assisted instructional program was evaluated with a group of students from Chemistry 302, the first-year college chemistry course at The University of Texas at Austin. Particular attention was paid to evaluating the instructional program in terms of its own objectives and expected outcomes, the latter being derived from the systems approach that was used in the development.

The principal emphasis of the study was not on demonstrating that computer-assisted instruction can be effective in the content area that was chosen but rather on investigating the usefulness and effectiveness of the particular development processes and design techniques that were employed. Thus, while the results are directly applicable only to a single specific program, they also illustrate in a more general way the feasibility of combining instructional guidelines and explicit decision rules with classroom-oriented content. It is felt that this approach provides the kind of empirical evidence that is a necessary part of instructional research.

PROGRAM DEVELOPMENT AND EVALUATION

A systems approach (Eraut, 1967) was used in developing the materials for the instructional program. The implementation of this approach has been described by several writers (Briggs et al., 1967; Glaser, 1965; Bunderson, 1971), all of whom include essentially the same procedures. The development process that was carried out as a part of the present study has, for clarity of presentation, been divided into the following steps:

(1) Specification of the terminal behavior and of the entering skills.

(2) Description of the intermediate objectives and arrangement of the various objectives into an appropriate learning hierarchy.

(3) Development of an instructional sequence.

(4) Design and preparation of the instructional materials.

(5) Developmental evaluation and revision based upon the results obtained thereby.

(6) Evaluation of the completed program.

This separation of the development process into several components does not, of course, imply that the activities were carried out independently of each other. On the contrary, there was considerable interaction among the various parts as development proceeded.

Intermediate design products developed in the process of preparing the instructional program were formulated and explicitly stated at appropriate points. These were concerned with sequencing, with individualizing mechanisms, and with the structure of the materials themselves.

Terminal Behavior and Entering Skills

The first step in developing the instructional materials required the specification of the terminal behavior and of the entering skills. The terminal behavior, of course, describes what the student should be able to do at the conclusion of the instructional program. It follows rather directly from the more inclusive goal of science education that students should be able to apply scientific principles to specific situations when a definite set of principles and some particular kinds of situations are introduced. Thus, the terminal behavior describes student abilities to apply the concepts and principles of chemical equilibrium to problems involving aqueous solutions of slightly soluble strong electrolytes or aqueous solutions of soluble weak electrolytes. Table 1 contains the set of related objectives and the minimum performance criterion that, taken together, describe the terminal behavior.

The entering skills describe certain abilities that the target population is expected to possess. They also serve to set a lower limit to the learning hierarchy, for it should not be necessary to instruct the student in those skills that he is presumed already to possess. The general target population for this program consists of first-year college chemistry students who are planning to major in one of the sciences or in engineering; more specifically, it is made up of students who are enrolled in Chemistry 302 at The University of Texas at Austin. It was assumed that the typical member of this population would have had, as mathematical preparation, at least two years of high school algebra and that he would also have either successfully completed one semester of college chemistry or demonstrated comparable achievement through advanced placement. It is readily apparent that such students are likely to possess a rather large number and variety of skills. Fortunately, it is not necessary to include all of these as entering skills but only those that form an explicit part of the learning hierarchy. It now becomes obvious that the learning hierarchy and the list of entering skills are closely related to each other. The two were actually developed simultaneously, and they have been separated here merely for convenience in discussing them. The entering skills that were identified as relevant to the other objectives of the instructional program are listed in Table 2. These entering skills, like the terminal behavior, have been described by means of behavioral objectives. It is seen that only two of the entering skills are related to chemistry, the other three being mathematical in nature.

Intermediate Objectives and Learning Hierarchy

In the second phase of the development process those intermediate objectives that need to be attained by the learner in order for him to progress from the entering skills to the terminal behavior were identified and defined in operational terms. The identification was accomplished by asking, in the manner first suggested by Gagné (1962a):

> What learnings, concepts, and skills will the student need in order to be able to perform the calculations required by this part of the terminal behavior?

The instructional designer, in attempting to answer this question, utilized both logical analysis and prior experience in teaching these specific applications. Table 3 has statements of the intermediate objectives.

Table 1

TERMINAL BEHAVIOR

At the conclusion of the instructional program, the student will be able to apply the concepts and principles of dynamic equilibrium to problems involving aqueous solutions of slightly soluble strong electrolytes or involving aqueous solutions of soluble weak electrolytes. The student will demonstrate his ability by obtaining numerical solutions to problems of the following types:

T-1a. Given the molar solubility of a strong electrolyte in distilled water or the concentration(s) of one or both of the dissolved ions in a saturated solution of a strong electrolyte in distilled water, the student will be able to calculate the solubility product of the electrolyte.

T-1b. Given the solubility product of a slightly soluble strong electrolyte, the student will be able to calculate the molar concentrations of each ionic species in a saturated solution of the electrolyte in distilled water.

T-2. Given the solubility product of a slightly soluble strong electrolyte and the molar concentration of one of the ions derived from the electrolyte, the student will be able to calculate the concentration of the other ion.

T-3a. Given the dissociation constant and the nominal concentration of a weak acid (or base) in distilled water, the student will be able to calculate any or all of the following:

(1) The concentrations of all ionic species.
(2) The pH and/or the pOH of the solution.

T-3b. Given the pH and the nominal concentration of a weak acid (or base) in distilled water, the student will be able to calculate the dissociation constant for the weak acid (or base).

T-4a. Given the molar concentrations of acid (or base) and of salt in a buffer solution and the dissociation constant of the weak acid (or base), the student will be able to calculate the pH of the buffer.

T-4b. Given the pH of a buffer and the dissociation constant of the weak acid (or base), the student will be able to calculate the relative concentrations of acid (or base) and salt.

T-4c. Given the pH of a buffer, the ionization constant of the weak acid (or base), and either the concentration of the acid or the concentration of the salt, the student will be able to calculate the remaining concentration (of acid/base or salt).

The student's ability will be evaluated by means of a posttest at the conclusion of the instructional program. He should be able to obtain correct solutions for 70% of the problems representing a sampling of the above types.

Table 2

ENTERING SKILLS

If he is to attain the terminal behavior, it is necessary that the student be able to demonstrate certain skills before entering the program. These are specified by the following prerequisite objectives:

P-1. The student will be able to write and interpret a chemical equation describing a given solution or ionization process. The student will demonstrate his ability in one or both of the following ways:

(1) Given the formula for a slightly soluble or soluble strong electrolyte or for a soluble weak electrolyte, the student will write a balanced chemical equation to represent the solution or ionization process.

(2) For a given amount (or concentration) of dissolved or ionized electrolyte, the student will be able to state the amount (or concentration) of each ionic species formed and/or the total amount (or concentration) of ions in solution.

P-2. The student will be able to define and use the term "molar" to describe the concentration of a solution. The student will demonstrate his ability in one or more of the following ways:

(1) The student will state in writing or will identify from among a group of written statements the definition of "molar."

(2) Given the name or formula for a substance and suitable information for determining its formula mass, the student will describe the contents of or a suitable method for preparing a solution of the given substance having any given molarity.

P-3. The students, with the help of a table of common logarithms, will be able to transform any given number, N, to $\log N$ and any given logarithm, $\log N$, to N.

P-4. The student will be able to generate and manipulate numbers expressed in exponential notation. He will demonstrate his ability in any or all of the following ways:

(1) Given a decimal number, the student will write the equivalent form of it in standard scientific notation.

(2) Given a number in exponential form, the student will write the equivalent decimal number.

(3) Given an arithmetic problem involving the operations of addition, subtraction, multiplication, division, raising to a power, and/or extracting a root in which some or all of the numbers are expressed in exponential notation, the student will solve the problem and write the solution in exponential form.

P-5. The student will be able to solve first- and second-degree equations of the following types:

$$ax + by + c = 0 \quad \text{and} \quad ax^2 + by^2 + cx + dy + e = 0$$

He will demonstrate his ability in one or both of the following ways:

(1) Given an equation of the above type, the student will solve it for either variable.

(2) Given an equation of the above type and information from which numerical values for the constants and for one of the variables may be deduced, the student will calculate the numerical value of the other variable.

The student should be able to answer correctly 70% of a group of questions and problems representing a sampling of the above objectives.

Table 3

INTERMEDIATE OBJECTIVES

I-1. The student will be able to define the equilibrium constant mathematically. He will demonstrate his ability by writing or identifying the equilibrium expression for the general gaseous reaction:

$$aA + bB \rightleftharpoons cC + dD$$

I-2. The student will have a functional understanding of the concepts and principles associated with chemical equilibrium and the equilibrium constant. He will demonstrate his ability in one or more of the following ways:

(1) Given the equation for a specific equilibrium reaction, the student will identify those substances whose concentrations appear in the equilibrium constant.

(2) Given the equation for a specific equilibrium reaction, the student will write the expression for the equilibrium constant.

I-3. Given the equation for a specific equilibrium reaction, the student will describe qualitatively the effect that a change in conditions, such as concentration, temperature, and so on, will have on the position of the equilibrium. He will also identify or state in writing the principle that is being used together with its name.

I-4. The student will state in writing or identify the definition of pH.

-5. The student can use the concept of pH and will demonstrate his ability in one or both of the following ways:

(1) Given the hydrogen-ion concentration of a solution, the student, with the help of a table of common logarithms, will calculate the pH of the solution.

(2) Given the pH of a solution, the student, with the help of a table of common logarithms, will calculate the hydrogen-ion concentration of the solution.

I-6. The student will be able to write the expression for the ion-product for water, to specify the numerical value of this constant at room temperature, and to calculate the hydrogen-ion concentration, the hydroxyl-ion concentration, the pH, and/or the pOH when given any one of the preceding quantities.

The entering skills, the intermediate objectives, and the several objectives included in the terminal behavior were then joined together into a learning hierarchy. The learning hierarchy shows diagramatically the relationships between the various objectives and indicates which objectives are considered to be prerequisite for the mastery of others. Here again both logical analysis and empirical observations derived from teaching experience were relied upon. Figure 1 presents the learning hierarchy. The letters and numbers that are used refer to the various objectives as listed in Tables 1, 2, and 3.

Instructional Sequence and Individualizing Mechanisms

The next step in developing the instructional materials required the translation of the learning hierarchy into an instructional sequence. Since the present state of instructional design theory does not offer any overall rationale for this operation, common sense, intuition, and the logic of the learning hierarchy itself provided most of the directions. The following generalizations guided the planning of the instructional sequence:

(1) Those terminal objectives, or component parts of the terminal behavior, that are most closely related to each other should be placed in nearest proximity in the sequence.

(2) Related terminal objectives should be presented in what appears to the designer to be the order of their increasing complexity.

(3) The various intermediate objectives should be introduced at the point where they will be most immediately useful to the learner.

(4) Materials not related to either the intermediate or the terminal objectives should be placed in the instructional sequence at those points where the student first has need for them. Such materials would include, for example, information about the program itself and about the conventional forms to be used in constructed responses as well as short reviews of selected entering skills.

(5) Opportunities should be provided for the student to practice the terminal objectives not only singly but also in related groups (Gagné, 1970).

Unless the instructional program is to be completely linear with all students using exactly the same materials in precisely the same order, the instructional sequence itself should be a variable one. Individualizing mechanisms make it possible for students to have different pathways through a set of instructional materials. These mechanisms may be entirely under the control of the learner; i.e., the student may be allowed to choose his own pathway through the instructional program. They may, on the other hand,

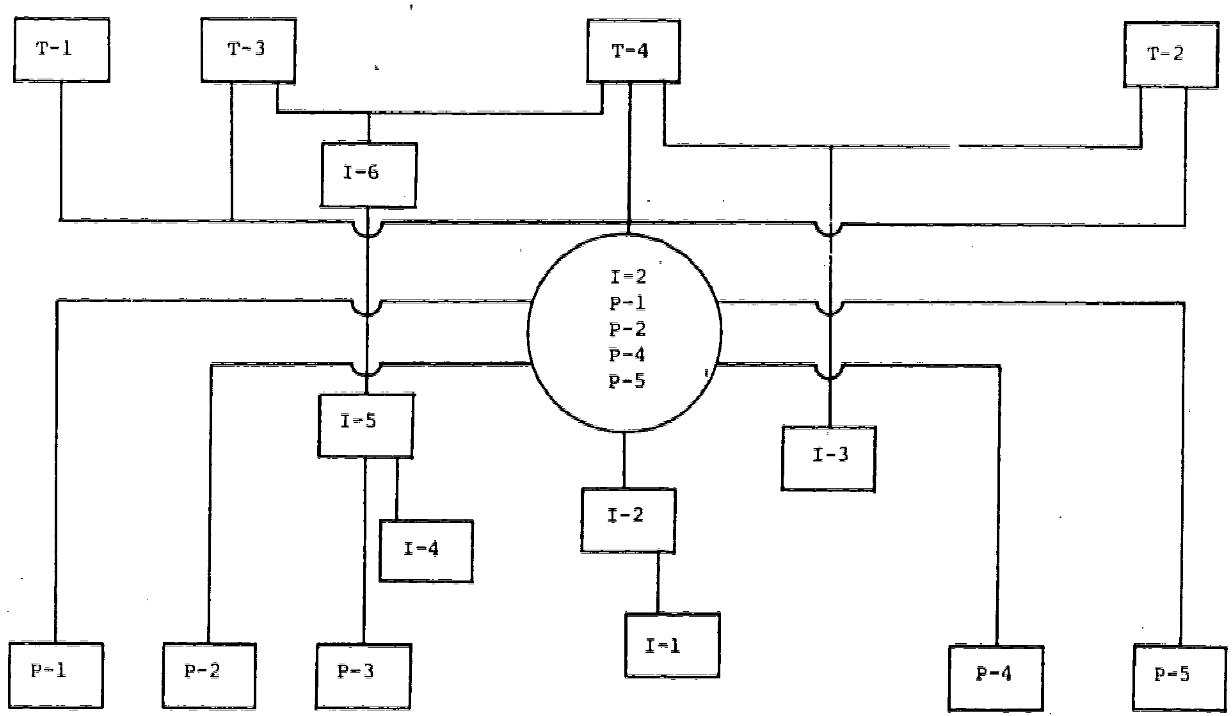

FIGURE 1.--Learning Hierarchy.

be wholly under program control, in which case the program itself contains rules and directions for defining the individual student's movement through the instructional materials. It is also possible, as was done in the present program, to combine the two forms of control, leaving some decisions to the student and reserving others to the program itself.

For those individualizing mechanisms that were retained under the control of the program, the following decision rules were used:

(1) A tutorial sequence that is related to an intermediate objective is to be bypassed if the student can demonstrate adequate mastery of the principles and concepts involved. The demonstration of "adequate mastery" is defined as responding correctly to at least 85% of the criterion items presented prior to the tutorial sequence.

(2) A practice sequence that is related to a terminal objective is to be bypassed if the student can demonstrate adequate attainment of the behavior involved. The demonstration of "adequate attainment" is defined here as solving without error a problem of the same type in the corresponding tutorial sequence.

(3) A problem series within a practice sequence is to be terminated when the student can demonstrate adequate attainment of the behavior involved. Demonstrating "adequate attainment" is, in this case, defined as solving without error the preceding problem in the practice sequence.

(4) Where a practice sequence contains problems described by several related terminal objectives, the student should be presented only problems that are similar to the items he missed on the immediately preceding mastery test.

These decision rules served the dual purpose of insuring that the student mastered one objective before moving to another one and of relieving him of additional instruction in an objective that he had already mastered. This emphasis upon mastery is in accordance with the theoretical position of Gagné (1962b) and with the empirical evidence accumulated by Silberman and Carter (1965).

For a computer-assisted instructional program, the learning sequence is best summarized by a flowchart showing the orderly progression through the major components of the program and identifying the principal points at which the decisions governing this progression are made. Figure 2 is a highly simplified and abbreviated form of such a flowchart for the present program. The central portion of this figure contains those parts of the instructional sequence that are used by all students, including both the tutorial materials related to the terminal objectives and also the various diagnostic and mastery tests. The tutorial segments that provide instruction in the intermediate objectives are located on the left side,

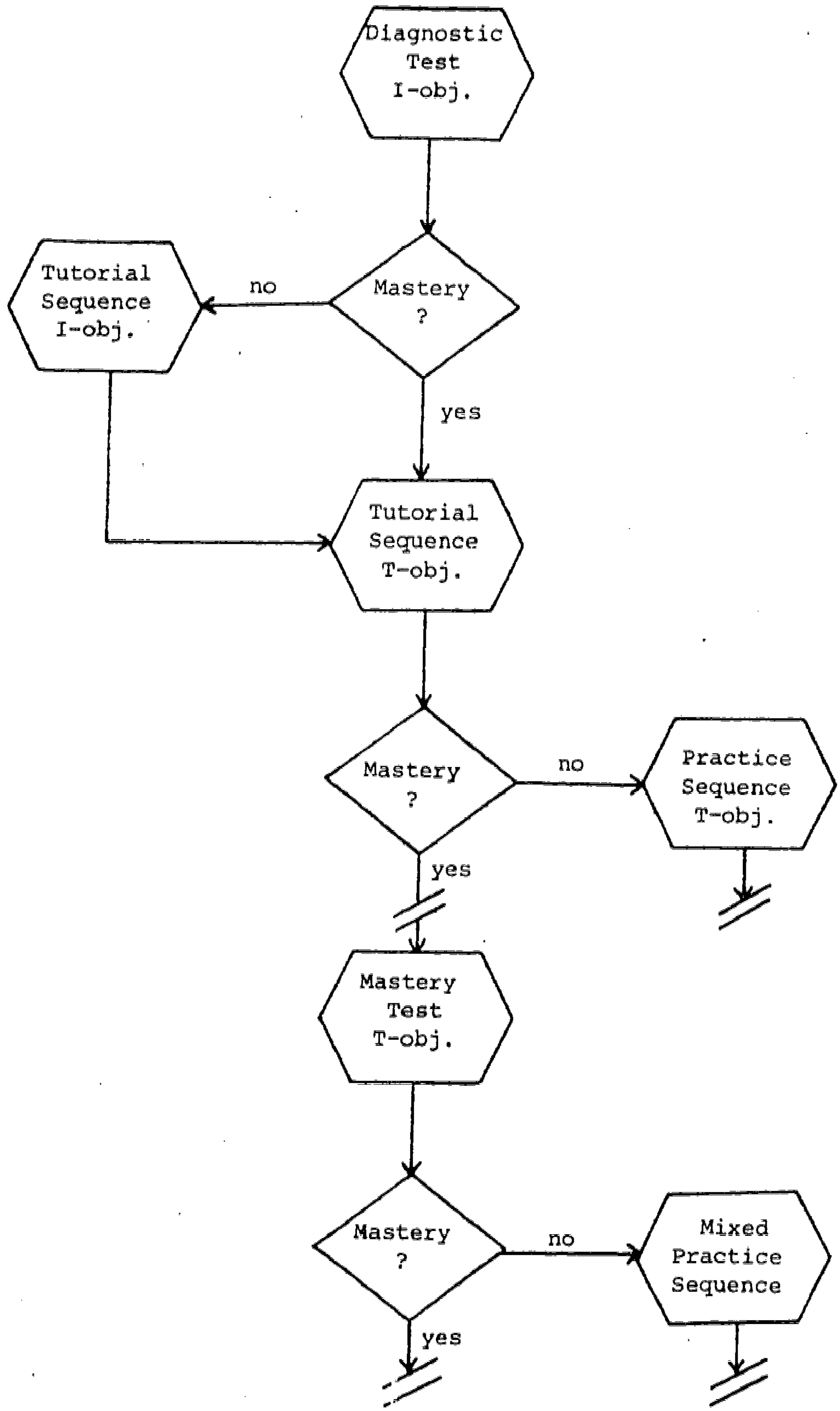

FIGURE 2.--Instructional Flow Chart.

and the groups of practice problems for the terminal objectives are on the right side. The total program contains two diagnostic tests and three tutorial sequences for the intermediate objectives, two mastery tests over groups of related terminal objectives and two practice sequences associated with these, and a final criterion test over the entire terminal behavior. There are also various auxiliary tutorial sequences related to the terminal objectives. These are accessed from the main tutorial sequences either at the option of the student or by the program itself where the student's responses indicate the need for supplementary instruction. Finally, the instructional program contains a few brief review sequences for some of the mathematical entering skills. The student decides for himself whether or not to make use of these.

The program is constructed so that the student is not permitted to go through any part of the instructional materials more than one time. Where normal operation of the decision rules would lead to the repetition of a sequence, the student is told to see the proctor for individual help.

Instructional Materials

The actual writing of the instructional materials can be considered as the fourth step in the development of the program. This was actually begun as soon as the general outline for the learning sequence had been decided upon, and the writing itself suggested changes in and provided detailed structure for the learning sequence. In addition, the information gained through the developmental evaluation of the early instructional segments guided not only the revision of these segments but also the initial construction of the later ones. There is, thus, much interaction between the processes of arranging the learning sequence, preparing the instructional materials, and observing student use of the various parts of the program.

Several generalizations, some of which were suggested by Gagné's *The Conditions of Learning* (1967), served as guidelines for writing the instructional materials:

(1) When a principle is to be learned, instruction should provide for recall of the relevant concepts.

(2) The student should be encouraged through verbal cues to organize and grasp the meaning of the principle for himself.

(3) When a principle is to be applied, instruction should provide for the recall of the principle, if such recall is necessary.

(4) The student should be encouraged to apply the principle for himself, and directions should indicate procedures without presenting solutions.

(5) Verbal cues, where possible, should be directed toward a specific learner difficulty.

(6) The instruction should incorporate both large and small steps. As long as the student is able to demonstrate progress, steps should be large; when the student encounters difficulty, steps should become smaller, and repetition should increase.

Since this program has been designed to perform a real teaching function with respect to the application of equilibrium principles, the instructional sequences pertaining to the terminal objectives are quite comprehensive. The student is expected not only to obtain the correct answer to the particular problem that has been presented to him, but also to analyze the general aspects of the equilibrium reaction under consideration, and so on. On the other hand, because it was anticipated that the student would receive additional instruction about the concepts and principles of chemical equilibrium from other sources, the materials selected for the intermediate objectives were severely limited in scope and only those aspects of immediate usefulness to the student were included. Even this restricted instruction, of course, was omitted for the student who could show that he had already mastered the necessary principles.

The instructional materials consist of questions to be answered and problems to be solved by the student, together with a minimum amount of expository text to be read by him. After the student has responded, he receives some kind of comment on his answer from the instructional program. This might be nothing more than mere confirmation in the case of a correct response. It was frequently possible to predict, on the basis of first-hand experience with real students, some of the incorrect responses that were likely to be made; thus, feedback information for the students could be made specifically relevant to these anticipated errors and the responses they would lead to. For other incorrect responses, verbal cues and hints of a more general nature were presented to the student. Since the process of applying knowledge was regarded as primary importance rather than the product of such application, the student was rarely simply told the correct answer after a single incorrect response. For much the same reason, sample problems and worked-out examples for the student to follow were deliberately omitted from the instructional materials.

In the tutorial and practice sequences of the instructional program, the student was usually required to construct his own response even when the question itself contained strong cueing. Because of this, the identification and incorporation of equivalent expressions for the correct answers and also for the anticipated incorrect ones became a necessary part of writing the instructional materials and one to which the developmental evaluation made important contributions. The items for the diagnostic and mastery tests, on the other hand, were all framed as multiple-choice questions.

The instructional program was coded in Coursewriter III for use with the IBM 360/50 system. The bulk of the material was displayed to the student by means of an IBM 2740 typewriter terminal; the student also used this typewriter keyboard for his responses. Those portions of the tutorial sequences that seemed awkward for typing, such as, for example, the mathematical expressions for equilibrium constants, were put on a precoded 16mm film strip. This technique was also used for the items making up the diagnostic and mastery tests. An IBM 1512 random access projector displayed individual frames from the filmstrip to the student as they were needed in the program.

Developmental Evaluation

Developmental evaluation, although it is being treated here as a separate step, proceeded concurrently with the preparation of the instructional materials and consisted of two kinds of activities. The purpose of the first of these was to examine the objectives and test items incorporated in the program; that of the second was to observe students using various parts of the instructional materials.

Examination of objectives and test items. The statements of objectives contained in the entering skills, in the intermediate objectives, and in the terminal behavior were evaluated by groups of science and mathematics educators. These individuals assessed the appropriateness of the various instructional objectives for the courses to which they had been related and the suitability of the test items as measures of student attainment of the objectives.

The test items for determining the student's mastery of the entering skills were used with two groups of Chemistry 302 students while the program was being written. Based on these preliminary observations, the test was revised to remove or simplify the most difficult items (those answered correctly by less than 40% of the students), and short reviews of some of the mathematical skills were added to the instructional program itself.

Individual observations of students. During the time that the instructional materials were being prepared, a total of about 20 individuals tried out various portions of the program, ranging in length from short sequences of tutorial questions up to essentially the entire program. Those who participated at this stage as students were not always members of the defined target population, but this diversity and the critical acumen of some of the advanced students are considered advantageous in developmental evaluation. Graduate and undergraduate students in science education and in chemistry were included, and the students were observed closely by the author as they interacted with the instructional materials.

Several kinds of revision stemmed from these observations. The simplest was the location and correction of errors that had been made in coding the material into the Coursewriter III language. Additional ways of expressing an acceptable answer were also identified and incorporated into the program. Similarly, new and previously unanticipated incorrect responses were inserted, and appropriate feedback was provided for them. Finally, those sequences that proved to be ineffective were completely rewritten, usually in an expanded form. In those cases where considerable revision was made in the materials, the new program was tried out again with another student or group of students.

Testing and Evaluation of the Completed Program

In order to obtain quantitative data with which to judge the effectiveness of the completed program, a group of students from one section of Chemistry 302 was selected to use the computer-assisted instructional materials during the spring semester, 1970. These students were also expected to attend the regular class meetings of the section. Consequently, the CAI program, while it served to supplement and reinforce the large-group instruction in the same problem-solving and application skills, did not have an independent instructional role within the total framework of the course.

Since it was felt that a meaningful evaluation of the program itself depended upon the students completing it without extensive reviewing, such as, for example, in preparation for an examination, a time schedule was arrived at in which the CAI program would be available for a period of six weeks immediately preceding an hour test that would contain similar kinds of problems. Because the program as developed forms a single cohesive unit, rather than a group of semi-independent modules, much of the needed data could not be obtained unless and until a student completed all of the program. As an incentive for them to do this, the cooperating professor responsible for this section granted credit for homework assignments to those students who finished instructional materials.

From a group of 145 students who volunteered to participate in the study, 35 were randomly selected. The remaining volunteers were used, where needed, as a comparison group. The following information was available for all of the students:

(1) Level of entering skills in mathematics and in chemistry as measured by the entering skills test prepared as part of the CAI program but given, in this case, separate from the program itself.

(2) Pre-program attainment of the terminal behavior as measured by test items prepared for the instruction program but used, again, separate from the program itself.

(3) Attitudes toward chemistry and toward CAI, both before and after the period during which the experimental group was using the CAI program, as measured with a semantic differential test developed by Castleberry (1969, p. 89) for a similar CAI study.

(4) Hour-test scores on program-related and non-program-related questions.

(5) Final examination scores on program-related and non-program-related questions.

In addition, for those students who completed all of the instructional program, measurements of the program path length and of the post-program terminal behavior were obtained. The program path length was defined, as the total number of critical labels in the instructional program as used by a particular student, the term "critical label" denoting those labels that contain a substantial ammount of cognitive content to which the student must respond in some way. The post-program terminal behavior was measured by a set of criterion items in the last part of the instructional program; these items were parallel with but not identical to the ones used to determine the pre-program attainment of the terminal behavior. Finally, the members of the experimental group responded to an informal, written questionnaire designed to elicit reactions to the total program.

R E S U L T S   O F   T H E   E V A L U A T I O N

This study has involved an admittedly small and ultimately self-selected student sample; consequently, it is subject to hazards both in interpretation and in generalization, and these are mentioned at appropriate points. These defects must not obscure the positive aspects. Computer-assisted instruction, especially as applied to educational research, is still a relatively new field; it is also undeniably an expensive form of instrumentation. It can, however, be used to explore otherwise inaccessible questions, and this exploration can lead to the discovery of unexpected effects. Although such effects need to be validated by replication and by further study on a larger scale, the initial effort is not without significance. In these circumstances stringent statistical tests may be less important than the mere fact that previously unconsidered phenomena have been given attention.

Program Evaluation

The criteria for evaluating the instructional program were contained in a set of statements of expected outcomes. It is felt that these outcomes, or others similar to them, are implied whenever systems analysis is used in the production of educational materials. In this case, however, they have been explicitly stated.

(1) It is expected that students will be able to demonstrate competence in the terminal behavior at the conclusion of the instructional program.

The mean pre-program and post-program attainment of the terminal behavior for the 14 students who completed the instructional program are shown in Table 4. The mean score on the program posttest is significantly higher ($F = 105$, $p = 0.0$) than the mean score on the pretest, the latter being adjusted to a basis of 12 items. In addition to the significant change in the mean score, each individual student who completed the instructional program achieved a higher score on the post-program test than he had on the parallel-item test given before the program.

Table 4

MEAN ATTAINMENT OF TERMINAL BEHAVIOR

	Mean	Maximum Possible	Range	Standard Deviation
Pre-program	1.1	6	0 - 2	0.5
Post-program	9.4	12	5 - 12	2.0

Table 5 summarizes the mean scores for the experimental and the comparison groups on the program-related and non-program-related portions of the third hour-test, which was given immediately after the period during which the members of the experimental group were using the CAI program. In addition to that for the entire experimental group, data are also included for three subgroups within this group: (1) those who completed the instructional program, (2) those who started the program but did not complete it, and (3) those who did not start the program. Table 6 contains the corresponding information for the final examination in the course.

The mean score on the program-related items is significantly higher than the mean score of the comparison group in the case of the hour examination; on the program-related part of the final examination the difference between the two groups is smaller and of questionable significance. When the non-program-related questions on these two tests are considered, the differences between the two groups are small and not statistically significant. This comparison of performance on program-related questions using the total experimental group probably tends to dilute the effect of the instructional program, since not all of those who were selected for the experimental group made equal use of the materials.

The results for the various subgroups within the experimental group indicate that those students who completed the instructional program performed better on program-related questions and problems, both on the hour test and on the final examination, than did the members of the other subgroups; on the other hand, their scores on the remaining items were not significantly different from those made by members of the experimental group who made less or no use of the instructional program. The extent to which these three groups of students should be compared is questionable, since they are obviously self-selected. It was not possible, however, to detect any differences among them with respect to mean scores on either standardized aptitude tests (SAT) or the preliminary tests used with this study (Lasater, 1971, pp. 148-149). The members of the three subgroups might also be expected to differ in factors such as motivation, efficiency of study habits, etc. While it was not possible to measure such traits directly, significant variation in them among the groups would be expected to yield differences in non-program-related course performance. Such differences were not observed with either the hour test or the final examination scores used in this study.

Table 5

THIRD-HOUR EXAMINATION SCORES

Descriptive Statistics	Description of Groups					F Ratio for Group Differences			
	Experimental				Comparison				
	Total	Subgroups[a]							
		A	B	C		Total/Comp	A/B	B/C	A/C
Program-Related Items (Maximum = 36)									
Mean	28.2				21.5	9.15***			
Range	0-36				0-36				
SD	9.7				10.5				
N	26				96				
Mean		32.4	22.8	25.0			6.38**	<1	1.70*
Range		20-36	0-34	12-32					
SD		4.9	12.5	11.3					
N		14	9	3					
Non-Program-Related Items (Maximum = 64)									
Mean	32.4				30.5	<1			
Range	10-57				2-64				
SD	14.4				15.3				
N	26				96				
Mean		31.9	34.3	28.6			<1	<1	<1
Range		9-60	6-59	22-36					
N		14	9	3					

[a]Subgroups: A = Completed the instructional program.
B = Started the instructional program.
C = Did not start the instructional program.

*$p > .20$ (two-tailed test)
**$p < .02$ (two-tailed test)
***$p < .002$ (one-tailed test)

Table 6

FINAL EXAMINATION SCORES

Descriptive Statistics	Description of Groups: Experimental: Total	Experimental: Subgroups[a]: A	Experimental: Subgroups[a]: B	Experimental: Subgroups[a]: C	Comparison	F Ratio for Group Differences: Total/Comp	A/B	B/C	A/C
Program-Related Items (Maximum = 36)									
Mean	24.8				21.8	1.97**			
Range	6-36				0-36				
SD	9.9				10.7				
N	29				106				
Mean		30.9	20.6	16.4			9.51***	<1	11.9***
Range		18-36	6-36	10-31					
SD		6.4	9.8	8.5					
N		14	10	5					
Non-Program-Related Items (Maximum = 139)									
Mean	93.6				91.0	<1			
Range	47-132				24-139				
SD	22.4				26.5				
N	29				106				
Mean		97.7	93.2	82.4			<1	<1	1.66*
Range		55-129	47-132	65-113					
N		14	10	5					

[a]Subgroups: A = Completed the instructional program.
B = Started the instructional program.
C = Did not start the instructional program.

*$p > .20$ (two-tailed test)
**$p < .02$ (two-tailed test)
***$p < .002$ (one-tailed test)

(2) It is hoped that all students who complete the program will attain the minimum level of competency specified in the terminal behavior.

Of the 35 students who were originally selected to use the computer-assisted instructional materials, 14 completed the entire program; this corresponds to 40% of the total starting group. If only those students who were members of the experimental group and who were still active in the course at the end of the semester, i.e., who took the final examination, are considered, the 48% of the CAI group finished the instructional program.

The minimum criterion specified in the terminal behavior states that the student "should be able to obtain correct solutions for 70% of the problems representing a sampling of the above types." Since the post-program criterion test contained 12 items, a score of 9 was considered adequate. Of the 14 students who completed the program, 10 scored 9 or more on the criterion test; this represents 71.4%.

(3) It is expected that a program having many branching points based on response history will yield a substantial amount of individualization when used by diverse students.

The mean path length for students completing the CAI program was 44 "critical labels" with a standard deviation of 5 labels. Figure 3 shows the distribution of observed path lengths. The program design itself sets a lower limit of 11 labels for the path length and permits, within the existing decision rules for branching, a theoretical maximum of about 85 labels although it was not anticipated that any student would actually use a program of that length. The observed range of path lengths is from 36 to 53, the lower limit of this range being well above the minimum and the upper limit being considerably below the theoretical maximum; thus, the shortest program path length contained about two-thirds as many labels as the longest.

(4) It is assumed that an ability to demonstrate the possession of certain entering skills is a necessary prerequisite for completing the instructional program and mastering the terminal behavior.

The student's ability to demonstrate the possession of certain prerequisite skills was measured by a written test, and the minimum criterion set for these entering skills was influenced more by the realities of what members of the target population could actually do than by any arguments about what they ought to be able to do. Since the minimum criterion specified 70% correct responses, a score of 7 or more on the 9-item entering skills test was required as a demonstration of competency. Twenty-two of the 35 students in the experimental group obtained such a score. Those falling below the minimum were allowed to take the program anyway because an assessment of the effect of the student's ability to demonstrate these skills on his performance with the instructional program was considered an important part of the evaluation.

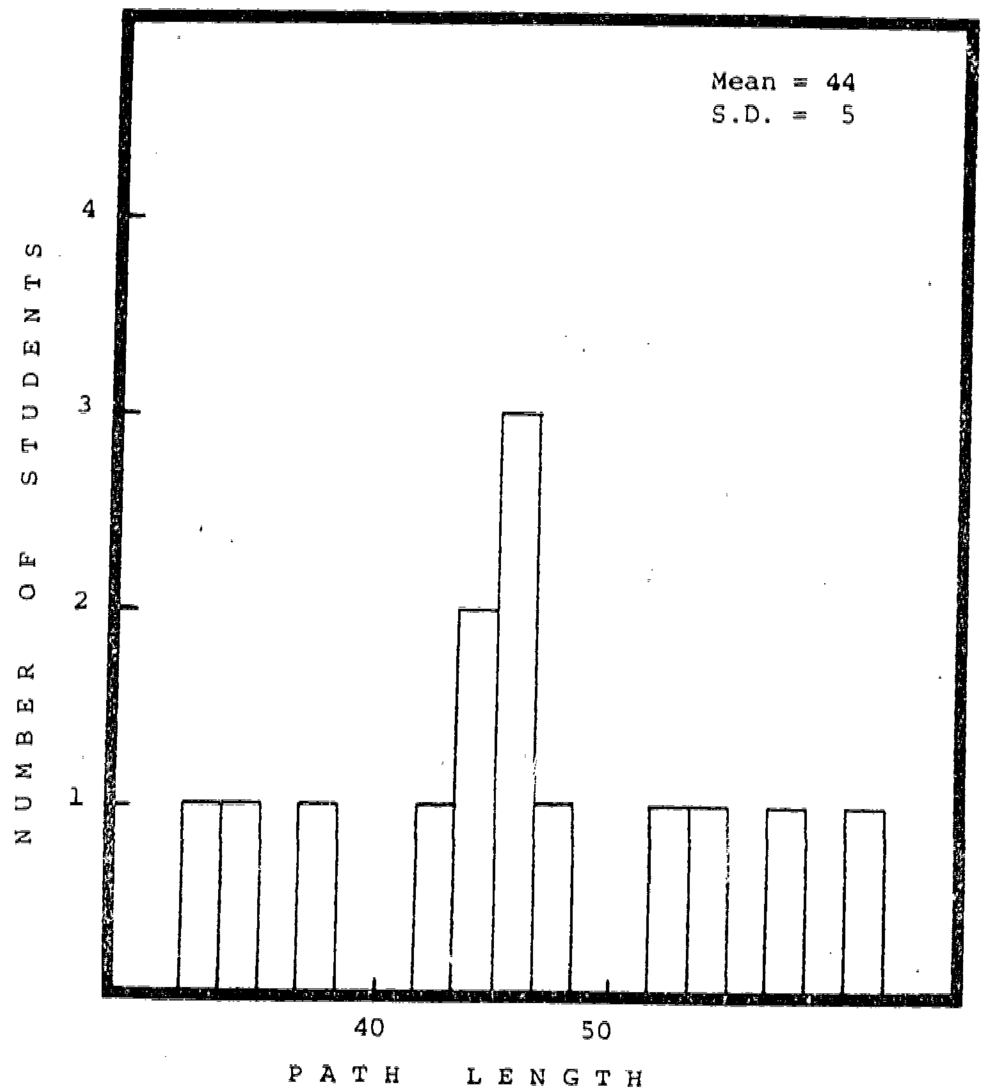

FIGURE 3.--Path-Length Distribution for Students Completing the Instructional Program.

The mean path lengths and the mean post-program test scores for those students who could and those who could not demonstrate competency in the entering skills are given in the upper part of Table 7. It is apparent that, *for those students who actually completed the entire instructional program*, the mean path length for those who demonstrated competency in the entering skills is not significantly less than the mean path length for those who did not demonstrate this competency. The mean posttest score was only very slightly higher for the former group than for the latter.

Since observations made by the experimenter while the students were using the instructional program had indicated that they encountered more difficulty with the mathematics than with the chemistry, it seemed worthwhile to investigate the mean path lengths and the mean post-program test scores for those students who could and could not demonstrate competency in the mathematics section of the entering skills test. The same criterion of 70% correct responses was used; thus, competency was defined as a score of 5 or more on the six mathematical items of the entering skills tests. The lower part of Table 7 presents the results when students are grouped according to mathematical competency. It should be noted that while the groups shown in the two parts of this table contain the same numbers of students they do not contain the same students. Both the difference in mean path length and the difference in mean posttest score have become magnified by the change in grouping, and the former has now become statistically significant.

In Figure 4 the observed program path lengths are plotted against the measured level of entering skills for the 14 students who completed the instructional program. The equation for the least-squares straight line through the data points, which is graphed on the same coordinates, is included; the quantity R, which is also shown, can be interpreted as a measure of the "goodness of fit" of the linear equation to the observed points. The graph of Figure 4 suggests that a relationship does exist between the student's path length and the level of his entering skills; however, the probability of obtaining this set of points in the absence of any dependency between the variables is large ($p = 0.225$ with a one-tailed test).

Because the path length appears to be more sensitive to the mathematical portion of the entering skills test than to the total score on this test, the observed program path lengths were plotted against the scores on that part of the entering skills test that directly concerned mathematical objectives. Figure 5 represents this graph. The supplementary information with the curve corresponds to that of Figure 4. The higher value of R, as compared with that of Figure 4, indicates a closer agreement between the actual path lengths and the values that would be predicted from the linear equation. The probability of obtaining the observed data points with unrelated variables ($p = 0.018$) has become low enough so that the result is considered significant, particularly in view of the small number of cases involved.

Table 7

MEAN VALUES OF PROGRAM VARIABLES

Variables	N	Mean	Range	S.D.	F Ratio
Minimum Level of Entering Skills					
Path Length					
With	11	43.5	36-51	4.0	<1
Without	3	46.3	37-53	8.3	
Program Posttest					
With	11	9.5	7-12	1.6	<1
Without	3	9.0	5-12	3.6	
Minimum Level of Entering Skills in Mathematics					
Path Length					
With	11	42.5	36-51	4.1	8.51**
Without	3	50.0	48-53	2.6	
Program Posttest					
With	11	9.7	7-12	1.8	1.83*
Without	3	8.0	5-10	2.6	

*$p > .20$ (two-tailed test)
**$p < .02$ (one-tailed test)

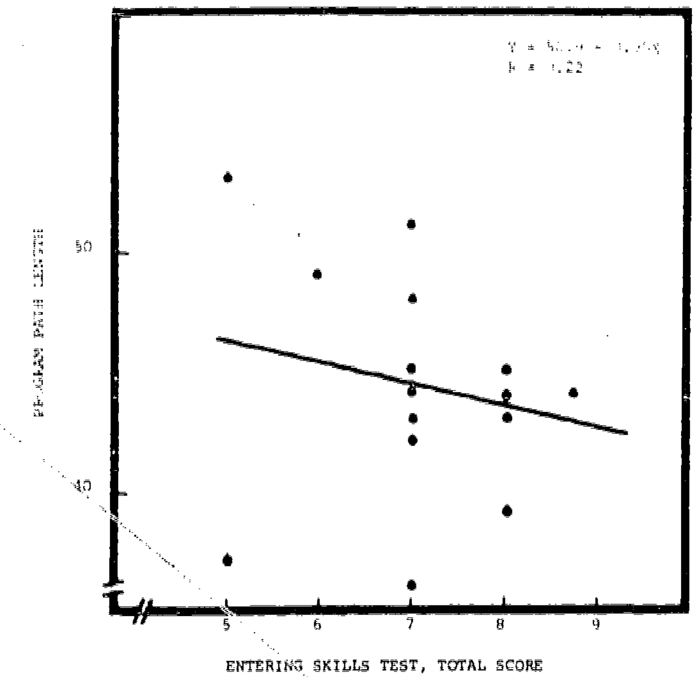

FIGURE 4.--Program Path Length as a Function of the Level of Entering Skills.

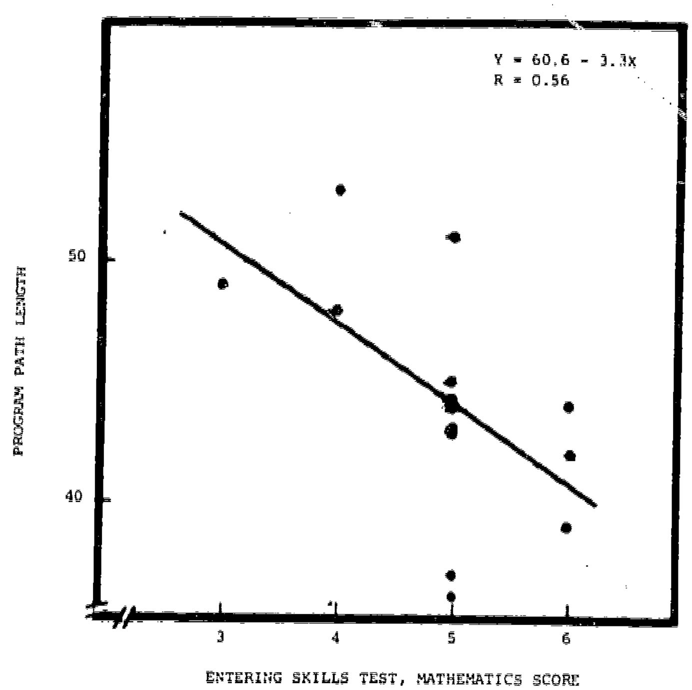

FIGURE 5.--Program Path Length as a Function of the Level of Entering Skills in Mathematics.

The post-program test scores are plotted in Figure 6 against the measured level of entering skills for the students completing the entire program. Here, also, the supplementary information is similar to that of Figures 4 and 5. Comparison of Figure 6 with Figure 4 shows that there is apparently much less relationship between the post-program test score and the level of entering skills than between the path length and the level of entering skills.

The results that have been presented and discussed up to this point seem to indicate that the student's ability to demonstrate competency in the entering skills does not greatly influence his success with the instructional program, particularly as estimated by his score on the criterion test at the conclusion of the program. These results, however, pertain only to those students who completed the materials, since it was only for them that the necessary data could be obtained. If all of the students who were selected to use the CAI program are considered, then it is possible to examine whether competency in the entering skills seems to affect the student's likelihood of completing the mater'ais. This information is compiled in Table 8. Although the differences between the two groups of students are not statistically significant, they do suggest a trend. The student who demonstrates competency in the entering skills seems somewhat more likely both to begin the instructional program and, having begun it, to see it through to completion. At the same time, the student who does not demonstrate competency in the entering skills, if he succeeds in completing the materials, seems to have about an equally good chance of mastering the terminal behavior.

Effects on Student Attitudes

Although the instructional program was not directly planned to influence student attitudes toward either the content or the instructional method, attitude measures were taken as a part of the evaluation. A semantic differential test was used to assess attitudes toward chemistry and toward the use of the computer as an instructional aid; the same test was administered to the entire section before and after the period during which the experimental group was using the instructional program. The results that were obtained are summarized in Tables 9 and 10, the former containing scores on that part of the test that was designed to measure attitudes toward the computer and the latter the scores for attitudes toward chemistry.

It is seen from the data in Table 9 that the mean attitudes of both the CAI group and the comparison group toward the use of the computer as a study aid declined slightly during the experimental period, the decline of the CAI group being somewhat less than that of the comparison group. The attitudes of the two intact groups are not, however, significantly different for either administration of the test. The results for the three subgroups within the experimental group, as summarized in the lower part of Table 9, show that the students who finished the program had somewhat more positive attitudes than did those who started it but did not finish. Interestingly enough, those students who made no use of the instructional materials had significantly less positive attitudes about CAI at the end of the experimental period than did those who had actually used the program.

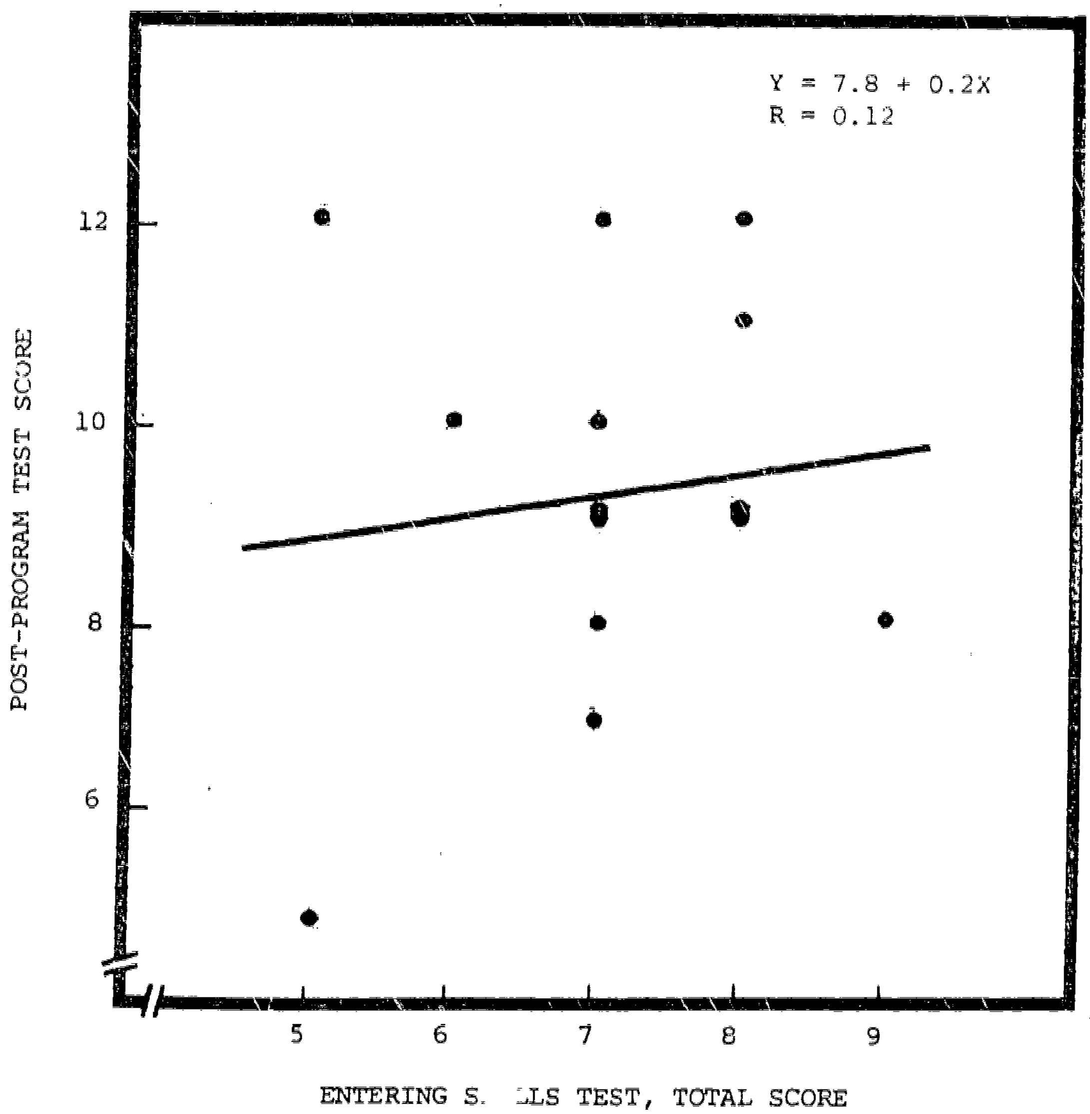

FIGURE 6.--Post-Program Score as a Function of the Level of Entering Skills.

Table 8

USE MADE OF CAI PROGRAM BY STUDENTS WITH AND WITHOUT MINIMUM LEVEL OF ENTERING SKILLS[a]

Use Made	With				Without			
	N	%Total	%Start	%Complete	N	%Total	%Start	%Complete
Started Instructional Program	18	82			7	54		
Completed Program	11	50	61		3	23	43	
Attained Criterion	8			73	2			67
Did Not Attain Criterion	3			27	1			33
Did Not Complete Program	7	32	39		4	31	57	
Did Not Start Instructional Program	4	18			6	46		
TOTAL N	22				13			

[a]Decimal points omitted.

The mean scores for the various groups and subgroups on the part of the test designed to measure attitudes toward chemistry are given in Table 10. Again there is a slight decline in the average attitude of both the experimental group and the control group between the first and second testing; however, as before, there is no evidence that the two groups differ significantly from each other. The results for the three subgroups within the experimental group show no significant differences among them although the attitudes of those who made maximum use of the program are again the most positive and those who made no use of the program show the least positive attitudes.

Relationship Between Program Variables

If all students were to demonstrate equal mastery of the terminal behavior at the conclusion of instruction, there would obviously be no relationship between the student's mastery and any other instructional variable. Since the students who completed the present instructional program did not all demonstrate the same degree of mastery, it seemed interesting to ask whether attainment of the terminal behavior (as measured by the post-program test) was related to the amount of instruction the student received (as measured by his path length through the program). In Figure 7 the post-program test scores for the students who completed the instructional materials are plotted as the dependent variable against the students' path lengths through the program; the least-squares line through the points has been graphed on the same coordinates. The equation for this line and the value of R associated with it and the data points are also included. The probability of obtaining these results with variables that are actually unrelated is only 0.0436 with a two-tailed test. This suggests the existence of an empirical relationship between the two program variables.

Since the results have already indicated that the path length is probably a function of the initial ability of the student to demonstrate competency in the mathematical entering skills, a further reasonable question was whether the post-program test scores are related to the path lengths for students who have the same level of entering skills in mathematics. This was approached by assuming that the criterion test score is linearly related to both the path length and the level of entering skills in mathematics. The prediction equation that is obtained from the measured data under this assumption is:

$$S = 25.7 - 0.3L - 0.7M$$

where:

S = post-program test score on terminal behavior.

L = path length.

M = score on mathematical part of the entering skills test

and where the coefficients have been computed by the method of least squares. If, for a given level of entering skills, the post-program test score and the path length are unrelated, then the value of the coefficient of L would be expected to be zero. The probability of obtaining the calculated value under these conditions if fairly small ($p = 0.0339$ with a two-tailed test), indicating that the path length and the post-program score continue to be inversely related when the level of mathematical entering skills remains constant.

Table 9

ATTITUDES TOWARD THE COMPUTER AS A STUDY AID

Descriptive Statistics	Description of Groups: Experimental Total	Experimental Subgroups[a] A	B	C	Comparison	*F* Ratio for Group Differences: Total/Comp	A/B	B/C	A/C
Pre-Program Scores (Maximum = 70)[b]									
Mean	56.5				57.7	<1			
Range	38-70				39-70				
SD	7.6				7.4				
N	35				114				
Mean		55.2	56.1	58.6			<1	[illegible]	1.13*
Range		38-69	45-70	49-70					
SD		8.1	7.4	7.3					
N		14	11	10					
Post-Program Scores (Maximum = 70)b									
Mean	54.2				51.9	<1			
Range	35-67				10-70				
SD	8.8				12.1				
N	27				80				
Mean		57.2	54.9	45.0			<1	5.10**	8.76**
Range		49-67	39-67	35-55					
SD		5.3	10.0	9.1					
N		13	9	5					

[a]Subgroups: A = Completed the instructional program.
B = Started the instructional program.
C = Did not start the instructional program.

[b]This score indicates a highly positive attitude.

*$p > .20$ (two-tailed test)
**$p < .02$ (two-tailed test)
***$p < .002$ (one-tailed test)

Table 10

ATTITUDES TOWARD CHEMISTRY

Descriptive Statistics	Description of Groups					F Ratio for Group Differences			
	Experimental				Comparison				
	Total	Subgroups[a]							
		A	B	C		Total/Comp	A/B	B/C	A/C
Pre-Program Scores (Maximum = 70)[b]									
Mean	53.0				53.5	(No statistically significant differences were found)			
Range	29-70				31-70				
SD	10.0				8.5				
N	35				114				
Mean		51.3	52.9	55.4					
Range		29-70	37-63	41-69					
SD		11.4	9.6	8.8					
N		14	11	10					
Post-Program Scores (Maximum = 70)[b]									
Mean	49.4				48.6	(No statistically significant differences were found)			
Range	22-68				22-69				
SD	12.4				10.8				
N	27				80				
Mean		51.6	48.8	44.8					
Range		30-68	22-65	36-54					
SD		11.7	15.8	6.5					
N		13	9	5					

[a]Subgroups: A = Completed the instructional program.
B = Started the instructional program.
C = Did not start the instructional program.

[b]This score indicates a highly positive attitude.

*$p > .20$ (two-tailed test)

**$p < .02$ (two-tailed test)

***$p < .002$ (one-tailed test)

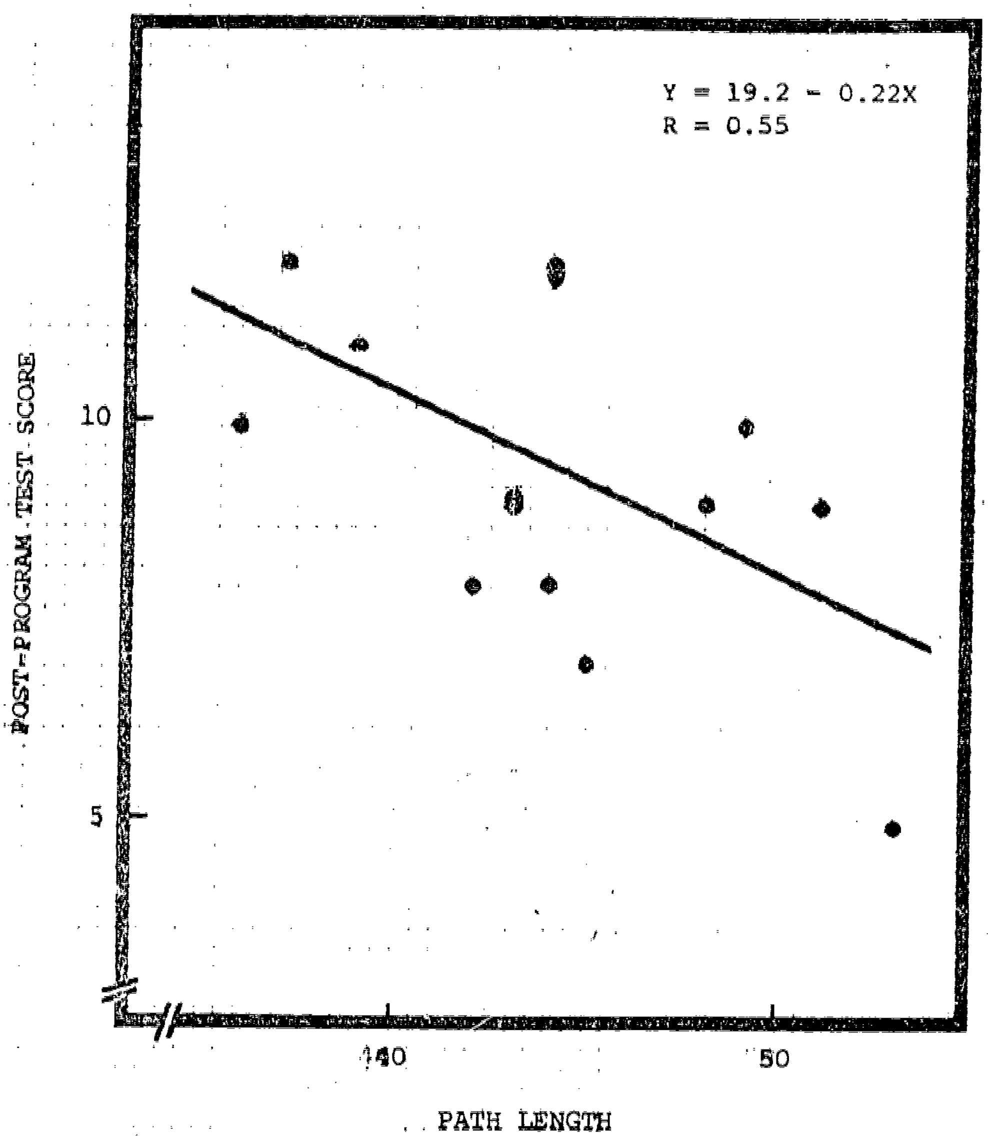

FIGURE 7.--Post-program Test Score as a Function of Program Path Length.

CONCLUSIONS

Program Development

The procedures that were used for developing the program and for designing the instructional materials appear to be relatively successful. The evidence, both from the developmental evaluation and from the formal testing, indicates that students do indeed acquire additional skill in the terminal behavior; moreover, the explicit decision rules incorporated in the program provide variables that can be changed in specifiable ways for instructional research.

Program Evaluation

The principal emphasis of the testing phase was directed toward assessing the effectiveness of the instructional program in terms of its own expected outcomes. Consequently, several of the conclusions are directly linked to these same expected outcomes. Some, however, are related to other aspects of program usage.

Conclusions relating to expected outcomes. These conclusions are based on observations made with a student sample that was, in effect, doubly self-selected. The students not only volunteered to participate in the study but also chose to complete the instructional program. The possible biases introduced by this are considered in the section on Limitations.

The expected outcomes that served as evaluative criteria are listed in the Results of the Evaluation section of this report. It was concluded that the program satisfied two of the expected outcomes (numbers 1 and 3) but fell short of meeting a third one (number 2). In the case of the final expected outcome (number 4), the results proved to be only partially congruent.

(1) The results of the study indicate that students who complete the instructional program demonstrated increased competence in the terminal behavior at the conclusion of it and that this increased competence was maintained over at least a short period of time.

(2) Approximately 70% of the students who completed the instructional materials attained the minimum criterion of 70% correct responses as measured by the program posttest.

(3) It was concluded, on the basis of the variance in the path-length distribution, that the program provided a considerable range of individualization when used by a variety of students.

(4) The results suggest that, at least under the conditions of this study, the student who is lacking in the entering skills is probably less likely to complete the instructional program. On the other hand, for those students who do finish the materials, the initial level of entering skill does not seem to greatly affect either the post-program test score or the likelihood of reaching the minimum criterion of performance. The student whose deficiency is in the mathematical component of the entering skills would, however, probably follow a longer path length and thus use more of the instructional material.

Conclusions relating to student characteristics and program variables. The program path length and the post-program test score are considered to be the important program variables. It was not possible to show any relationship between either of these variables and standardized aptitude measures (Lasater, 1971, pp. 136-138.

(1) The student's expected path length through the instructional materials is inversely related to his demonstrated competence in the mathematical entering skills as measured prior to the start of the program.

(2) There appears to be an inverse relationship between the student's path length through the program (treated as the independent variable) and his post-program test score (treated as the dependent variable), and this relationship persists when allowance has been made for different levels of entering skills in mathematics.

Conclusions relating to student attitudes. Although not a primary part of the study, the attitude measures did yield some interesting results.

(1) It appears that the instructional program might have had a positive effect on the attitudes toward CAI of those students who made maximum use of it.

(2) For those who volunteered but did not make use of the program, however, a "sour grapes" attitude was noted.

DISCUSSION

Limitations on the Study

Several factors serve as possible limitations on the generalizability of the conclusions based on the results of this study. In the first place, the students who participated were all volunteers. Also, since they were expected to attend lectures in which the same kinds of applications were discussed, their motivation to complete the instructional program was largely intrinsic. Thus, any extension of the results to a general population of students, especially when coupled with the extrinsic motivation that would result from a definite instructional role being assigned to the CAI program is valid only to the extent that one is willing to assume that the behavior of these students is typical of the larger group. It is the opinion in this report that those results reflecting primarily cognitive aspects of the student's behavior, such as the achievement of the terminal behavior, the effects of entering skills, etc., would probably remain at least qualitatively correct. Where attitudinal or motivational effects are important, as, for example, in the fraction of students completing the program, there would seem to be considerably less certainty about any extrapolation.

A second limitation concerns the amount of influence the instructional program had in improving student performance with respect to the terminal behavior. That such performance was better at the end of the program than it had been before the program was started seems undoubtedly true. On the other hand, it cannot be assumed that all, or any identifiable part, of this learning was the direct result of the program itself, since the students were simultaneously exposed to instruction in the same skills during the class lectures. Even if it is conceded, as is warranted by the evidence, that the students who completed the CAI program performed at a higher level than their classmates, it cannot be assumed that this "excess" skill could not have been acquired by them in some other way, such as by working more problems from the textbook on their own.

Implications of the Study

Instruction and learning. The results indicate that students with less competency in the mathematical entering skills identified as specifically necessary for the program content followed a longer path length than did students with greater competency. This is in agreement with Gagné's generalization: "The course of learning for any individual is importantly affected by the capabilities he brings to the instructional situation" (1965). That the post-program performance on the criterion test was apparently not related to the level of these entering skills even though the program itself provided no specific instruction in them suggests two possibilities. The student may have been relearning through practice certain capabilities that he had previously possessed and thus, in effect, have been improving his proficiency in the entering skills as he progressed through the program. It is also possible that the items on the criterion test, being multiple choice rather than of the constructed response form used in most of the instructional program, placed less demand on the student's mathematical skills; the student who had made a calculation error might well not find his answer among the listed choices and in that case would be led to recheck his arithmetic. It should be noted again that relatively few of the students with a low level of entering skill completed the instructional program.

The results also suggest that a student's criterion score is related inversely to his path length through the instructional program, and this dependence seems to persist even when allowance is made for differing levels of entering skills. This effect can be accounted for by assuming the existence of one or more abilities that, when present to a larger degree, permit the student to master the instructional material more rapidly and with better comprehension. It is tempting to speculate that in this case the relevant ability might be one that facilitates the translation of verbal material into mathematical statements. If this hypothesis that the relationship between path length and post-program test score is due to the influence of differing student abilities of some kind is correct, then the results of the present study are in agreement with the finding of Schurdak (1965) that "more able" students achieved higher criterion scores with less exposure to the course material.

Of the 35 students originally selected to use computer-assisted instruction, only about 40% completed the entire program. This observation is comparable to Castleberry's results in a similar kind of testing experiment with first-year chemistry students; he reports (1969, p. 87) that of a CAI sample of 100 students chosen from among a larger group of volunteers, 38 were "fully participating" and 36 had "limited participation." There is no real reason for assuming, however, that this same "drop-out" rate would persist if an instructional program had any sort of "required" status within a total course structure. There are a variety of possible explanations for the observed completion rate. Students who lacked a satisfactory level of entering skills may have simply found the material too difficult; the kind of instruction, as described by the instructional guidelines, may have been inappropriate for some of the students. There may have been a lack of sufficient motivation, since failure to complete the program was not penalized in any way. Finally, the students may have simply become discouraged by the delays and frustrations associated with malfunctioning of the computer system. The students themselves, when asked about this as a part of the evaluation, mentioned most frequently the demands on their time made by other courses or by other parts of the chemistry course.

The results obtained with the attitude inventories would seem to have some important implications, especially for any widespread or routine use of CAI. Although it seems clear that experience can affect student attitudes about CAI, it is not obvious whether the influence comes from the instruction itself, from the subject-matter content, from the computer used to manage the instruction, or from some combination of these. The results seem to indicate that, at least in this study, more than one factor must have been important. Among those students who used some portion of the instructional program, the ones who completed all of it had slightly more positive attitudes about CAI at the end of the semester than did those who did not complete the materials; however, the differences were not large. On the other hand, those members of the experimental group who did not even start the instructional program developed attitudes that were significantly less positive than the ones expressed by students who had used the materials. The source of these negative feelings is not at all clear. These students may have been influenced by the poor performance of the computer system if, for example, they arrived for a scheduled appointment to find that CAI was not operating; or they may have been affected indirectly by information about the program from other students, but this seems unlikely in view of the generally more positive attitudes expressed by the others; or their attitude change may have been the result of factors entirely extraneous to the CAI program itself, a possibility suggested by the simultaneous presence of less positive attitudes toward the subject matter itself.

Instructional design. It is the opinion in this paper that this program was, considering the constraints within which it was used, reasonably successful in meeting its instructional objectives. There remained, in spite of several revisions, one tutorial sequence with which students still had considerable difficulty, and this should be rewritten before further use is made of the program. In addition, the results of the study provide an empirical basis for revising and clarifying two of the expected outcomes.

The preceding comments do not, of course, imply that there is any reason for regarding this as the "best" kind of instruction for these terminal objectives or for this target population. It is possible that one or more changes in the instructional guidelines or decision rules might improve the efficiency of learning, as measured either by increased student competency in the terminal behavior or by decreased program length, for students using the program. Such changes, if they are to contribute useful information for future instructional designers, must be explicit and must be validated by empirical evidence.

Improvement in the student's ability to apply concepts and principles to specific situations is an important goal of most science education. The general structure of the program developed for this study, including the instructional guidelines and the decision rules for branching, would seem to require only minor modifications in order to be suitable for many terminal behaviors involving applications of concepts and principles, not only in chemistry but also in other areas of science.

Computer-assisted instruction. The opinion expressed in this report, based on the experience reported, is that computer-assisted instruction provides very valuable instrumentation for instructional research. Both the requirements for its effective use and the advantages it can offer have been adequately described in the literature (Zinn, 1967; Gerard, 1967); however, the large investments in time and money that are required will probably continue to limit its utilization in this area.

The various difficulties that were encountered with the computer system and their possible effects on the student users and on the results of this study emphasize again the importance of operational dependability. Although this has, of course, been recognized by others (Gentile, 1967; Suppes, 1966), the unfortunate consequences with regard to the collection and interpretation of research data are rarely mentioned.

Suggestions for Further Study

The results of this study suggest several questions that might be worth investigating, and these are listed below. The first four topics involve this particular CAI program; the other two are general areas under which related problems have been grouped.

(1) What is the effect on post-program scores and on program path lengths of changing the decision rules to allow students to decide for themselves whether and for how long they want to solve practice problems? Does this change have any effect on the percentage of students completing the program or on student attitudes toward CAI? A study of these questions has been initiated, and the results will appear in a subsequent report.

(2) What are the use characteristics, particularly in terms of student completion rates but also including program path lengths, post-program test scores, and the relationships between these program variables and the measured competencies of the students using the materials, of the program when it is integrated into the total course structure for those students who utilize it? How do these characteristics compare with the results of the present study?

(3) What is the effect of providing remedial instruction for students who cannot demonstrate competency in the mathematical entering skills? Where students do not receive remedial instruction, does their ability to demonstrate competency in these skills change between the beginning and the end of the instructional program?

(4) Is the inverse relationship between program path length and post-program test score attributable to differences in innate ability among the students? Can this ability be identified and measured?

(5) Very little is known about student attitudes toward instructional methods. What factors are most influential in shaping student attitudes upon initial exposure to an instructional method? What is the effect of these initial attitudes on later experiences of the student with the same or a highly similar instructional method? How can negative attitudes toward an instructional method be altered?

(6) There seems to be a need within instructional research for groups of studies in which instructional variables are changed one at a time in specific and systematic ways. These are the kinds of studies that should provide the empirical data needed to improve instructional theory. With the present study as a starting point, various directions might be taken. For what kinds of "application" objectives can these instructional guidelines and decision rules be used? Are they appropriate for secondary school students? Are they useful with even younger learners? If the instructional guidelines and decision rules are not suitable for a variety of objectives or for learners of different ages, can they be modified and made more general?

R E F E R E N C E S

Briggs, L. J., Campéau, P. L., Gagné, R. M., & May, M. A. *Instructional media: a procedure for the design of multi-media instruction, a critical review of research, and suggestions for future research.* Pittsburgh: American Institutes for Research, 1967. Pp. 28-52.

Bunderson, C. V. The computer and instructional design. In W. H. Holtzman (Ed.) *Computer-assisted instruction, testing, and guidance.* New York: Harper & Row, 1971.

Castleberry, S. J. The development and evaluation of computer assisted instruction programs on selected topics in introductory college chemistry. Unpublished doctoral dissertation, The University of Texas at Austin, Austin, Texas, 1969.

Eraut, M. R. An instructional systems approach to course development. *Audiovisual Communication Review*, 1967, 15, 92-101.

Gagné, R. M. The acquisition of knowledge. *Psychological Review*, 1962, 69, 355-365.

Gagné, R. M. Military training and principles of learning. *American Psychologist*, 1962, 17, 83-91.

Gagné, R. M. The analysis of instructional objectives for the design of instruction. In R. Glaser (Ed.) *Teaching machines and programmed learning, II. Data and directions.* Washington, D.C.: National Education Association, 1965. Pp. 21-61.

Gagné, R. M. *The conditions of learning.* New York: Holt, Rinehart, & Winston, 1967. Pp. 33-59.

Gagné, R. M. Some new views of learning and instruction. *Phi Delta Kappan*, 1970, 51, 468-472.

Gentile, J. R. The first generation of computer assisted instructional systems: an evaluative review. *Audiovisual Communication Review*, 1967, 15, 23-53.

Gerard, R. W. (Ed.). *Computers and education.* New York: McGraw-Hill, 1967. P. 15.

Glaser, R. Toward a behavioral science base for instructional design. In R. Glaser (Ed.) *Teaching machines and programmed learning, II. Data and directions.* Washington, D. C.: National Education Association, 1965. Pp. 771-806.

Lasater, M. B. The development and evaluation of a computer-assisted instructional program involving applications of selected chemical principles. Unpublished doctoral dissertation, The University of Texas at Austin, Austin, Texas, 1971.

Schurdak, J. J. *An approach to the use of computers in the instructional process and an evaluation.* Yorktown Heights, N.Y.: IBM Watson Research Center, 1965.

Silberman, H. F., & Carter, L. D. The systems approach: technology and the school. In E. H. Goodman (Ed.) *Automated Education Handbook.* Detroit: Automated Education Center, 1965. Pp. IC 26-40.

Suppes, P. The uses of the computer in education. *Scientific American,* 1966, 215, 206-220.

Zinn, K. L. Computer technology for teaching and research on instruction. *Review of Educational Research,* 1967, 37, 618-634.