# The use of problem solving skills in computer-aided instruction: an evaluation

By: M.J. Safrit, Catherine D. Ennis & F.J. Nagle

Safrit, M.J., Ennis, C.D., & Nagle, F.J. (1988). The use of problem solving skills in computer aided instruction: An evaluation. *Journal of Computer Assisted Learning*, *4(4)*, 227-243.



**Abstract:**
When computer programs are designed to elicit problem solving skills by the user, this basic premise should be tested in an objective manner. The Health and Fitness Assessment program was evaluated as an interactive program with a substantial emphasis on the problem solving process. A method of analysis known as protocol analysis was used to demonstrate that 72% of the interpretive statements made by users verified the use of higher level mental functions to interact with the computer. Other evaluative data aided in improving the design of the program.
**Keywords:** Computers; Computer problem solving; Simulation; Protocol analysis; Software evaluation; Computer program evaluation.

**Article:**
As Johnston (1987) has aptly noted, the successful use of microcomputers in education is not merely a function of the availability of hardware. Simply purchasing expensive microcomputers and peripheral equipment will not bring about meaningful educational change. A plethora of other factors affect the quality of educational gains that potentially may be attained, not in the least the quality of software available for educational purposes. Many educators have decried the lack of good quality educational software (e.g. Adams & Jones, 1983; Preece & Jones, 1985). Johnston (1987) correctly identified the major problem in developing software for computer-aided instruction — the failure to specify what the program should do. These considerations include 'theoretical assumptions about the nature of learning, about student behaviour, motivations and reactions, and about the curriculum itself' (Johnston, 1987, p.41).

This has major implications for the evaluation of computer software designed to aid student learning. Other aspects of the program (e.g. screen display, user friendliness, terminology) should not be overlooked, but it is more important to evaluate the program as an educational tool rather than a technological tool. The literature is replete with articles dealing with software evaluation. The importance of preliminary field testing of new programs has been stressed by some authors (e.g. Roblyer, 1985), and others have listed criteria for the evaluation of courseware (e.g. Cohen, 1983; Fetter, 1984). However, much of the instrumentation for software evaluation is designed to be used by teachers and curriculum specialists. These evaluation forms usually require the evaluator to identify features represented in the program. For example, if problem solving is perceived to be one of these features, it is checked on a form. Yet, this is not sufficient evidence that problem-solving skills are actually applied by students who use the program.

The deficiencies noted in educational software in general are exacerbated in the health and fitness area, where the scope of published software is limited and mostly of the tutorial and/or data interpretation variety. Even these programs are, for the most part, primitive and sometimes replete with errors. There has been a great demand for software for health and fitness assessment, not so much because of the educational need in schools, but rather due to consumer needs in private settings, such as fitness centres. Generally, these programs can be categorized under one of two types. One type is a program designed to analyse a client's health and fitness status in a fitness centre. For example, the Tenneco Corporate Fitness Center in Houston, Texas uses a bank of microcomputers to store and display physical fitness data and present an exercise prescription for the employee.

The computer is used to monitor the employee's exercise behaviour. The second type of program is designed to aid teachers in interpreting fitness data. Scores on fitness tests are converted to norms, usually percentiles, and other descriptive statistics are calculated. In some cases, an exercise prescription is provided for each student; however, these prescriptions tend to be quite general in nature. Both types of programs can lead to improvement in the management of classes and fitness centres, but they do not create a learning environment for the student who is being trained to provide instruction and leadership in health and fitness programmes. These students must learn to assess the appropriate physical and physiological parameters and use this information to prescribe an exercise programme for future students or clients.

The purpose of this study was to evaluate a computer program designed to elicit the use of problem-solving processes in assessing health and fitness parameters and developing exercise prescriptions. Not only is the program described in this paper a unique one, but the analytical approach used to evaluate the program as a means of implementing problem-solving behaviours is one that is infrequently observed in the literature dealing with software evaluation. The program is known as Health and Fitness Assessment *(HAFA)* * and was developed in the Measurement Laboratory at the University of Wisconsin, Madison. The analytical approach is protocol analysis, a method for analysing verbal data reports based on an information processing model.

## The health and fitness assessment (HAFA) Program

The HAFA program was designed to give students an opportunity to interpret fitness parameters for a subject and utilize this information to develop an exercise prescription. As the program was formulated and developed, it seemed apparent that students would be required to make complex decisions while using the program. However, the validity of these types of programs should not rest on assumption alone. Thus, this study was undertaken to provide objective evidence of the suitability of the program as a problem-solving tool.

### *Description of the HAFA program*

The HAFA program gives a student the option of interpreting data for either a hypothetical case study or a real subject. The variables are physiological and fitness parameters which can be used to describe the subject's fitness status. After analysing the data for a subject, the student is asked to develop an exercise prescription for this subject. The student is then asked to assume the subject has followed the prescription for six weeks. New data are presented for the student to analyse. The prescription is modified if warranted by the new data. Two types of feedback (Cohen, 1985) are used in this program. One type is knowledge of results, whether the student selects the right or wrong response to a problem. The second type is informational feedback, which allows the learner to correct an error by providing sufficient information. A Help menu can be used to review relevant tutorials and tables of norms.

A series of special files (Cohen, 1985) are available for the instructor of the course. In the HAFA program, this system consists of a summary of response time data for each student, a provision for creating a student or subject file, a provision for deleting a student or subject file, an option allowing the addition or deletion of a name to the list of acceptable student users, and a summary of the results of a survey students take upon completion of the simulation.

Development of the program began in August 1986. The IBM PC C Compiler was used along with Assembly language for some of the subroutines. A Toolkit program was used to assist in developing the graphics. The computers in the Measurement Laboratory were configured in a token ring network with an IBM PC-AT serving as a host computer.

### *Problem-solving component of the program*

The student is given two major problems to solve. The first problem is to assess the overall fitness status of the subject; the second, to generate an exercise prescription appropriate for the subject given his or her fitness status. The design of the simulation allows a student to move step by step through a pathway to solve the problem. The student may solve the problem by answering the question immediately or may elect to access tables of norms or tutorials for additional information. When the former option is taken, the student is operating within the goal pathway of the simulation program while the latter is characteristic of the problem materials pathway.

Providing students with a choice between goal and problem materials pathways is consistent with the two-process problem-solving approach proposed in the literature. Davis (1966, p.42) has suggested that 'the primary value of this approach is that empirical results in many areas of problem solving seem quite amenable to 'explanation' under the suggested two-process interpretation'. A comparison of characteristics of the goal and problem materials pathways is presented in Table 1. In the goal pathway (Benjafield, 1971; Reid, 1951) respondents answer the question directly without the benefit of additional information. They solve the problem covertly (Davis, 1966) using mental processes to assemble the needed information without visual or manual manipulation. The answer may be arrived at either explicitly or implicitly (Berry & Broadbent, 1987). Respondents who solve the problem explicitly are conscious of the rules or strategies that they used to determine the answer; while those who solve the problem implicitly are not consciously aware of their cognitive processing (Lewicki et al. 1987).

**Table 1.** Comparison of characteristics of goal and problem materials pathways.

| Goal pathway* | Problem materials pathway* |
|---|---|
| 1. Respondent answers directly without additional information | 1. Respondent selects to use instructional material to develop response |
| 2. 'Correct' response is known to respondent | 2. 'Correct' response is not known without experimentation |
| 3. Covert† — needed information is processed internally with no observable behaviours | 3. Overt† — problem-solving process is externalized and is observable |
| 4. Explicit‡ — some respondents are able to state rules or strategies used to solve problems<br>Implicit‡ — some respondents unaware of information synthesized or rules used to solve problem (tacit solution) | 4. Explicit‡ — respondents are able to state rules or strategies used to solve problems |

*Benjafield (1971).

†Davis (1966) used these terms to categorize two-process trial and error problem-solving originally described by Duncker (1945), Maier (1945) and Reid (1951).

‡As the study of cognitive processing has become more sophisticated, cognitive psychologists (e.g. Berry & Broadbent, 1987; Lewicki *et al.* 1987) have focused on the respondent's conscious and unconscious understandings described as implicit and explicit problem-solving strategies.

In the problem materials pathway (Benjafield, 1971; Maier, 1945), respondents choose to use materials, in this case norms tables and tutorials, to guide their decisions. For example, they solved the problem overtly (Davis, 1966) by visually comparing the student's raw score with the table values and noting the correspondence percentile. Respondents were able to describe explicitly the process they were using to solve the problem (Berry & Broadbent, 1987).

The simulation reported here was flexible in that respondents could choose either pathway for each of the 178 pathway decisions. Thus, they were able to choose their preferred problem-solving strategy depending on the perceived difficulty of the problem.

Regardless of the problem-solving strategy or pathway selected initially (goal or instructional aids), an incorrect response required the student to seek additional help from the instructional aids. The program automatically placed a norms table on the screen. Upon request, one or more tutorials could be viewed. Subsequent incorrect answers resulted in the provision of additional information for the student on the screen. A student who failed to answer correctly after four incorrect responses was given the correct response and then presented the next problem.

### Evaluation of *HAFA*

After nine months of programming and considerable informal assessment in the formative evaluation vein, a more formal approach to evaluation was used. The purpose of the evaluation was two-fold:

1. to improve the program design,

2. to verify the use of problem-solving skills by students.

Six undergraduate students participated in the study. It was necessary to restrict the sample size due to the extensiveness of the analysis of results necessary for this type of study. However, as Ericsson & Simon (1984) have noted, a large sample is not necessary for comprehensive evaluation, especially if a reasonable strategy is utilized for the selection of subjects. In this study, the students were carefully selected to represent a wide range of ability levels, as reflected in their GPAs and their grades in a required undergraduate measurement *course.* Both males and females were included in the sample. All had attended eight laboratory sessions on microcomputer usage as part of the measurement course.

Each student spent two sessions working with the computer simulation program. The first session began with a training session, where the general purpose of the simulation was explained, and documentation was given to each student. Then the protocol was described. Because one of the purposes of the study was to examine the extent to which problem-solving skills were required to use the program, continual responses from the student were essential.

The methodology used in this study is known as protocol analysis using verbal reports as data (Ericsson & Simon, 1984). This approach will be described in more detail in the next section of the paper. The student was instructed to talk continuously while interacting with the computer. Although it was made clear to the students that they should relate technical problems associated with the program, reporting the thinking underlying their problem solving was emphasized. Ericsson & Simon (1984) refer to this as a 'think aloud' procedure.

At the end of the training session, the student was asked to solve a word puzzle using the 'think aloud' protocol. The student was then seated at a computer and given a chart to be used in interpreting data. A tape recorder was used to record all verbalizations. Ericsson & Simon (1984) discussed the advantages of allowing the user to work alone while using the program. They reported the tendency of subjects in verbal report studies to talk to the investigator if one was present, rather than conveying independent responses. Therefore, the students in this study were taught to turn the tape recorder on and off, and were left to interact with the computer independently. However, one of the investigators or an assistant was available in an adjacent room in case the student needed assistance or the continual flow of verbalization ceased. The ventilation system in the laboratory prevented the hearing of specific words being spoken by the student; however, the student's voice could be heard in the adjacent room. If the student stopped speaking, he or she was reminded to 'think aloud' continuously.

### Protocol analysis: verbal reports as data

Verbal reports have been used in many types of research over the years. However, a subject's verbalization has sometimes been erroneously viewed as unsuitable scientific data. This perceived deficiency can be avoided by using protocol analysis, which allows an information processing model of the cognitive processes to be used to provide a basis for incoding verbal protocols in an explicit and objective manner (Ericsson & Simon, 1984). Although there are many variations of information processing models, a generic model — with components common to all models — is adequate for protocol analysis. When used to interpret verbal data, the model assumes that information is stored in several memories having different capacities and accessing characteristics.

The use of principles of human information processing when developing courseware is highly recommended by Jay (1983).

An underlying assumption of the model is that any verbal report of cognitive processes would be based on some form of the information stored in short-term or long-term memory. One of the two forms of verbal reports that most closely reflects the cognitive processes is the concurrent verbal report, which was used in this study. Verbal probes, where a subject is given a fixed set of alternative responses or asked to respond to specific questions, were not used. Ericsson & Simon note that the use of probes may produce reports not closely related to the thought process.

In implementing the 'think aloud' process, the student was asked to verbalize thoughts generated in the course of performing the simulation. Ericsson & Simon's second level of verbalization was stressed. This level involved an explanation of thought processes, which included the recoding of information in short-term memory and linking this information throughout the simulation. In other words, this level reflected interpretive and reasoning processes. Information at this level was used to verify the use of problem-solving skills.

### *Application of protocol analysis*

In analysing the transcripts of the subjects' tapes, the primary goal was to identify interpretive statements that could be tied to the cognitive processes used by the subject in utilizing the HAFA computer simulation. Initially, the tape was transcribed in double spaced non-paragraph form. The written transcript of the tape was pre-processed, which involved the segmentation of the verbal stream to allow the investigators to identify both relevant and irrelevant data. The pre-processed segments were put in protocol format, with each statement identified by the student's initials and the number of the statement. A portion of a transcript in protocol format is shown in Transcript 1. Each segment was assumed to constitute one instance of a general process. Cues used to identify segments were pauses, intonations, and indicators of the completion of a sentence. Then these segments were encoded into the terminology of the theoretical model. To accomplish this, interpretive statements were identified. Cues used to identify these statements were words such as if, so, and, because, and since. Whatever words or phrases followed the cue word were examined to determine whether they were used in an interpretive sense.

Two intercoder reliability coefficients were calculated using the scored-interval method (Hawkins & Dotson, 1975). The first coefficient reflected the agreement of two coders on the protocol statement number, while the second indicated their agreement on the category code for that numbered statement. The scored-interval method requires calculations of reliability in each of the coding categories. Randomly selected blocks of 100 to 350 statements were coded. Coefficients were calculated prior to the data coding and twice during the analysis process. Coefficients for the line number agreement ranged from 0.89 to 0.73. Category coefficients ranged from 1.0 to 0.50.

The protocol analysis produced five types of statement: transition, computer steps, interpretive, program, and read:

1. *Transition*: a shift from one section of the simulation to another;

2. *Computer steps*: student selection of an answer or a request for additional information (initiated by the simulation);

3. *Interpretive:* an explanation of the problem-solving process resulting in a decision made by the student;

4. *Program*: the identification of a problem associated with the program or an expression of confusion on the part of the student;

5. *Read*: the reading or paraphrasing of a case study or a tutorial.

Three types of statements — computer steps, interpretive, and program —were further analysed for the purpose of this evaluation. The students were also asked to use the 'think aloud' protocol while responding to a survey after the completion of the simulation. Their responses were then examined for interpretive statements.

**Transcript 1.** Portion of transcript for subject LM

| | |
|---|---|
| LM110 | Now I am moving to the 1-mile run. |
| LM111 | Each thing (screen) I move to is highlighted in a purple colour, and it draws my attention very easily. |
| LM112 | And I don't have any problems knowing exactly what I'm going to do. |
| LM113 | 8 minutes and 30 seconds for 1-mile run for a 5th grader seems quite fast to me. |
| LM114 | I'll go ahead and check the norms. |
| LM115 | Whoops it doesn't like it when I hit return. |
| LM116 | I don't follow directions all that well. |
| LM117 | Again I am going to the norms table which is Table 3. |
| LM118 | Choice A gives me norms by age for the 1-mile run. |
| LM119 | Age 11, again age is going across the top now. |
| LM120 | which is fine, which is different. |
| LM121 | 8:30 falls between, not as well as I thought actually, between the 60th and 70th percentile. |
| LM122 | In my mind it is almost, it's closer to the 60th percentile than the 70th. |
| LM123 | Now I'm facing cardiorespiratory fitness on the. . . . |
| LM124 | I am going to check the instructions I was given. |
| LM125 | It was close to the 60th percentile so it is going to be very close to average and above average. |
| LM126 | I think I have a better chance of. . . . |
| LM127 | I'm going back to the Help menu |
| LM128 | and back to the table. |
| LM129 | Pushing 3 and pushing E for 1-mile run. |
| LM130 | See if I can identify any closer my subject's time is 8:30 and I am going. . . . |
| LM131 | The difference between the two is 26 seconds — between the 60th and 70th percentiles. |
| LM132 | And if I add 13 seconds to the 70th percentile score that puts it a 8:25 which would be 65%. |
| LM133 | I am still 5 seconds higher than that. |
| LM134 | I think it is going to be above 61%. |
| LM135 | I am eyeballing this. |
| LM136 | If I had the calculator, I'd figure it out. |
| LM137 | I might know for sure whether I was going to get a right or wrong answer. |
| LM138 | I am going to say he is above average. |
| LM139 | But if he is, it is just barely. |
| LM140 | Got the correct answer. |
| LM141 | Every now and then this brain works alright. |

### *Identification of interpretive statements*

Constant comparison (Glaser & Strauss, 1967) was used to develop the coding system for the interpretive data. Constant comparison is a two-part inductive process that resulted, in this study, in a defined system of categorization. In the first part, 1268 interpretive statements from the six subjects were scanned to locate common terms or phrases. Similar statements were then grouped and rescanned to identify properties common to the group. The properties became the basis for the category definitions. The second part of the analysis consisted of the comparison of statements across categories to verify the integrity of the category membership, Statements were then enumerated. Efforts were made to establish categories that were mutually exclusive and to articulate definitions which directly reflected group properties. The categories emerged from the subjects' thought processes as reflected in protocol statements and were not imposed through an external categorization system. Examples of interpretive statements are shown in Table 2.

**Table 2.** Examples of interpretive statements

| | |
|---|---|
| LM27 | To evaluate his height, I need to see more norms for students his age. |
| FL265 | Now I am going to punch in 3, to see the normative table menu. |
| CC54 | I will look across until I find 40. |
| HS10 | 143 cm, calculated in inches, probably 2-70-2 143-70. |
| CC649 | For a 16-year-old, 11 minutes, 30 seconds is close to 11 minutes 49 seconds, closer to that than 11:08. |
| HS128 | I would say, frequency, minimal 3 times a week which is number 5. |
| HF156 | I would think he should develop (abdominal strength) because he should be above average. |
| LM85 | Also based on the distributions the largest percentage and the highest probability would be associated with normal. |
| HS486 | I guess when I picked average before, I meant that the measurement was below average in which she should be in body composition. I think the terminology is a little confusing. |

Interpretive statements were coded into seven major categories, listed below. Statements within each of the seven categories were further subdivided into subcategories. Definitions of the properties of subcategories were derived from the analysis.

(1) Understanding of simulation (U)

(a) Statements which acknowledge the subject's understanding of the computer program.

(b) Statements indicating the subject's understanding of the problem or the problem solving process. These include the student's understanding of problem/effective strategy, recognition of change in format of program (suggesting change in student's thinking process), or acknowledgement of need for change in student's problem solving strategy.

(c) Statements paraphrasing the information on the screen.

(2) Computer program (C)

These are statements related to the simulation itself.

(a) Student's problems with program (described in previous section).

(b) Statements related to steps in the simulation (described in previous section).

(3) Exclamation (EX)

Exclamations and editorial comments.

(4) Problem solving process (P)

Steps within the reasoning process used to derive an answer.

(a) Student's acknowledgement of need for resource information.

(b) Procedure for acquiring resource information.

(c) Procedure for locating data on table.

(d) Statement of relative proximity of given value to table values.

(e) Judgement — interpreting a given score; definite statement of the answer.

(f) Statements characterized by verification, correction, or clarification of earlier statement, hypothesis, or judgement.

(5) Rationale for selection (R)

(a) Rationale for choice based on previous information from course-work, readings, etc., not from the computer simulation (implicit knowing).

(b) Explicit statement of criteria or precondition used to make decisions.

(c) Maximum or minimum parameter values used to delimit choices before decision is made.

(d) Acknowledgement of several options as acceptable followed by a decision.

(e) Data from simulation used as a criteria or pre-condition.

(f) Information that was acknowledged previously; consciously stored for later retrieval.

(g) Understanding of the program expectations, procedures, etc.; used as rationale for decision.

(6) Questioning of program (Q)
(a) Questioning accuracy or rationale for correct answer identified in simulation.
(b) Expressing concern about limited options; being asked to select one option when the student knows that several options may be correct.

(7) Guessing (G)
(a) Student narrowed the answer to a small number of options, then guessed one response above or below previous choice.
(b) Random guessing; no statement of rationale.

## Results and discussion

Results indicated that students followed a systematic problem solving process leading to the development of an exercise prescription. The computer simulation provided boundaries for the decision making process through carefully structured problem solving. In the goal pathway, students made decisions based on the test score without additional assistance. If a correct response was given, the procedure took only one step in the simulation. It is clear from the protocol analysis, however, that students frequently used explicit problem solving strategies to synthesize information from practical experience, previous coursework, and an understanding of norms tables to arrive at a correct response.

The six students made a total of 178 pathway decisions during the assessment component of the simulation. These data are summarized in Table 3. Of these, 93 or 52% were selections favouring the goal pathway. Thirty-five per cent of these resulted in incorrect responses. The problem materials pathway was selected in 85 or 48% of the 178 possible choices.

**Table 3.** Frequency of pathway selection.

| | | Problem solving pathway | | | |
|---|---|---|---|---|---|
| | | Goal | | Problem Materials | |
| Student | No. of Responses | Selected | Incorrect | Selected | Incorrect |
| 1 | 30 | 29 | 13 | 1 | 0 |
| 2 | 30 | 4 | 0 | 26 | 8 |
| 3 | 30 | 9 | 3 | 21 | 6 |
| 4 | 30 | 28 | 9 | 2 | 0 |
| 5 | 28* | 17 | 7 | 11 | 8 |
| 6 | 30 | 6 | 1 | 24 | 5 |
| Total | 178 | 93 (52%) | 33 (35%)† | 85 (48%) | 27 (32%)‡ |

*Missing data.
†Percentage incorrect in goal pathway.
‡Percentage incorrect in problem materials pathway.

Thirty-two per cent were incorrect. Therefore, even though students selecting the problem materials pathway had access to table information, they made incorrect responses only slightly less frequently than students selecting the goal pathway.

When the data were analysed by pathway, there was evidence to suggest that students demonstrated pathway preferences. Three of the students exhibited a preference for the goal pathway. Data to support this claim are reported in Table 4. Eighty-four per cent of the combined responses from Students 1, 4 and 5 were made in favour of this pathway. When the frequency of incorrect responses were analysed by student, those preferring the goal pathway exhibited a 39% error rate (29 errors in 57 attempts). These same students, when selecting to use the problem materials (non-preferred) pathway, made 8 errors in 14 attempts, resulting in an error rate of 57%.

**Table 4.** Students preferring goal pathway.

| Student | No. of Responses | Problem solving pathway | | | |
|---|---|---|---|---|---|
| | | Goal | | Problem Materials | |
| | | Selected | Incorrect | Selected | Incorrect |
| 1 | 30 | 29 | 13 | 1 | 0 |
| 4 | 30 | 28 | 9 | 2 | 0 |
| 5 | 28* | 17 | 7 | 11 | 8 |
| Total | 88 | 74 (84%) | 29 (39%)† | 14 (16%) | 8 (57%)‡ |

*Missing data.
†Percentage incorrect in goal pathway.
‡Percentage incorrect in problem materials pathway.

The three other students in this study preferred the problem materials pathway. The results of this analysis are presented in Table 5. Students 2, 3 and 6 chose to make 79% of their selections in the problem materials pathway. When operating in their preferred problem-solving mode (problem materials), they exhibited 19 errors in 71 attempts for an error rate of 27%. The error rates for these students demonstrated little difference by pathway. They had a combined error rate in the goal pathway of 4 in 9 attempts for 27%.

**Table 5.** Students preferring problem materials pathway.

| Student | No. of Responses | Problem solving pathway | | | |
|---|---|---|---|---|---|
| | | Goal | | Problem Materials | |
| | | Selected | Incorrect | Selected | Incorrect |
| 2 | 30 | 4 | 0 | 26 | 8 |
| 3 | 30 | 9 | 3 | 21 | 6 |
| 6 | 30 | 6 | 1 | 24 | 5 |
| Total | 90 | 19 (21%) | 4 (21%)* | 71 (79%) | 19 (27%)† |

*Percentage incorrect in goal pathway.
†Percentage incorrect in problem materials pathway.

Although all students attempted to answer from the goal pathway at least four times, rarely were they able to demonstrate a *series* of correct answers. Student 1 attempted to answer 29 of the 30 questions using the goal pathway. However, only six consecutive questions were answered correctly. Students 4, 5, and 6 each completed a series of four consecutive correct answers. Students 2 and 3 answered only two consecutive questions correctly in the goal pathway. Therefore, although students may have preferred this mode of problem solving, they were generally unable to perform consistently. This was probably due to the technical nature of the information required to make an accurate assessment. Although the students had access to all necessary information, three chose not to utilize it 84% of the time. It is unlikely that students at this level of expertise possessed adequate working knowledge or experience to consistently arrive at an accurate response.

### *Interpretive statements*

The analysis of interpretive statements further demonstrated the extent to which the students used problem-solving skills in interacting with the computer. These statements can also be used to verify the relationship between an underlying information processing model and the use of the computer simulation program. In Table 6, the statements are summed according to code and section of the program across the three analyses. The statements are summed according to analysis across the seven codes in Table 7.

Three categories represent the understanding or use of problem-solving skills in this simulation. These are problem solving, rationale, and understanding. As shown in the last column of Table 6, over 40% of the interpretive statements denoted the use of problem-solving skills. Approximately 18% of these statements reflected the rationale students used to make decisions. Statements categorized as understanding the problem-solving process made up 14% of the total. The sum of the percentages in these three categories was 72%; thus, the identification of a substantial number of the interpretive statements provided evidence that students used problem-solving skills in interacting with the computer.

Most of the interpretative statements (58%) were made during the first case study analysis *(see* Table 7). Data from the analyses of problems, steps and interpretations were used as independent sources to confirm that the first analysis was the most difficult for students. The third analysis, in which students were required to input data, reflected an internal progression which increased the difficulty of the program over the case study analysis after a six-week exercise program, as revealed through the additional number of problems, steps, and interpretive statements. In the second analysis, a smaller percentage (13.7%) of interpretive statements were made; however, the percentage increased to almost 28% in the third analysis. This suggests that to maximize the use of information processing skills, instructors should pay special attention to the initial use of the simulation by students.

**Table 6.** Frequency and percentages of interpretive statements by category

| | Fitness parameters | | Exercise prescription | | Total | |
|---|---|---|---|---|---|---|
| Category | Freq.† | % | Freq.† | % | Freq. | % |
| Understanding* | 124 | 18.40 | 57 | 9.60 | 181 | 14.27‡ |
| Computer program | 78 | 11.57 | 82 | 13.8 | 160 | 12.62 |
| Exclamation | 34 | 5.04 | 43 | 7.23 | 77 | 6.73 |
| Problem* solving | 347 | 51.58 | 166 | 27.95 | 513 | 40.46‡ |
| Rationale* | 52 | 7.71 | 174 | 29.29 | 226 | 17.82‡ |
| Questioning | 17 | 2.52 | 29 | 4.88 | 46 | 3.63 |
| Guessing | 22 | 3.26 | 43 | 7.24 | 65 | |
| Total | 674 | | 594 | | 1268 | |

*Represent understanding or use of problem-solving skills in simulation.
†Number of interpretive statements summed across all three analyses.
‡Percentage of statements reflecting problem-solving; sum of 14.27%, 40.46%, and 17.82% = 72.55%.

Even a conservative interpretation of these results clearly points to the effectiveness of the program in requiring the use of higher level mental processes. Students had to analyse and synthesize information, and use the information to make decisions. At least 72% of the interpretive statements verified this conclusion. It would appear that the level of sophistication of the problem solving process could be increased by requiring the student to type words or phrases as responses rather than select responses from a series of options provided by the program. However, one of the specifications of the first phase of program development was to present the material in a format familiar to students. The provision of a set of choices seemed to meet this specification. Nonetheless, the results remain impressive in documenting the extensive use of problem solving skills throughout the simulation.

**Table 7.** Frequency and percentages of interpretive statements by analysis

| | Fitness parameters | | Exercise prescription | | Total | |
|---|---|---|---|---|---|---|
| Analysis | Freq. | % | Freq. | % | Freq. | % |
| 1* | 342 | 50.89 | 398 | 67.00 | 741 | 58.44 |
| 2† | 74 | 10.98 | 100 | 16.84 | 174 | 13.72 |
| 3‡ | 257 | 38.13 | 96 | 16.16 | 353 | 27.84 |
| Total | 674 | | 594 | | 1268 | |

*Hypothetical case study data prior to exercise programme.
†Hypothetical case study data after six-week exercise programme.
‡Actual case study data input by student.

### *Improvement of program design*

The verbal reports were also used to evaluate the program design. Statements reflecting students' perceptions of problems in the simulation program were coded and categorized.

Program statements were initially categorized by the portion of the program where they occurred:

**1** Analysis of fitness parameters;

**2** Development of exercise prescription;
**3** Transitions between segments;
**4** Survey.

The frequency of occurrence of these statements is summarized in Table 8. Most of the problem statements were made by the students during their first attempt to analyse a case study. In the first analysis, hypothetical case study data for a subject were examined prior to his or her participation in an exercise program. The number of program statements for this analysis are shown in the second column of Table 8. The largest number were those expressing Confusion (41) and Program problems (47). By the time the student completed the analysis of the second portion of the case study, the number of Confusion statements were reduced from 41 to 11 and the Program statements from 47 to 5. This suggested that many of their concerns were resolved by the time the program was used a second time. Greater familiarity with the technology or problem-solving process may have been a factor. It is also possible that the students simply overlooked a Program problem when it occurred the second time or worked around it. The next most frequent Program problems were identified as the lack of Efficiency (E) and Question (Q), with 22 and 24 statements respectively in the first analysis. Apparently the students' perceptions of these problems were tempered in subsequent sessions as their frequencies were reduced considerably. The third largest problem categories were Terminology and Typographical, with 10 and 8 statements identified in the first analysis. This information was summarized and given to the programmers to be used in improving the design of the program.

**Table 8.** Problem statements tallied by code.

| Code§ | 1st analysis* | 2nd analysis† | 3rd analysis‡ | Total |
|---|---|---|---|---|
| Efficiency (E) | 22 | 4 | 4 | 30 |
| Confusion (C) | 41 | 11 | 12 | 64 |
| Question (Q) | 24 | 5 | 3 | 32 |
| Program (P) | 47 | 5 | 18 | 70 |
| Terminology (TE) | 10 | 2 | 1 | 13 |
| Typographical error (TY) | 8 | 1 | 4 | 13 |
| Total | 152 | 28 | 42 | 222 |

*Analysis of hypothetical case study data prior to exercise programme.
†Analysis of hypothetical case study data after six weeks in programme.
‡Analysis of actual case study data input by students.
§E = suggestions to make program more effective.
C = confusion (problem originating with *student's* misunderstanding of computer program).
Q = question or disagreement.
P = problem originating with computer program.
TE = specific suggestion or comment about awkward or unclear terminology.
TY = typographical error, usually spelling.

The number of problem statements varied between students. One student in particular verbalized many more of these statements than the other students. Several problems, especially Terminology and Typographical error, were identified by more than one student. However, the statements were not tallied by the specific problem within a category, because the total number of statements was of greater interest.

## Future research

Subsequent investigations of the HAFA program will continue to focus on the problem solving processes used by students. Salomon & Gardner (1986) cited evidence of the value of holistic research paradigms when evaluating courseware, especially during the early phases of research. It would be interesting to compare expert and novice performances, using procedures described by Larkin et al. (1980). Perhaps the learning process could be made more efficient by exploring the kinds of processes an expert uses when solving problems. Ultimately, the most important goal in computer-based instruction is to teach students to make good decisions about significant problems in an efficient manner.

## References


Adams, A. & Jones, E. (1983) Teaching Humanities in the Micro-electronic Age. Open University Press, Milton Keynes.

Benjafield J. (1971) Evidence of two-process theory of problem-solving. *Psycho*nomic Science, 23(6), 397-399.

Berry, D.C. & Broadbent, D.E. (1987) The combination of explicit and implicit learning processes in task control. Psychological Research, 49(1), 7-15.

Cohen, V.B. (1983) Computer courseware: Development and evaluation. Educational Technology, 23(1), 9 - 14.

Cohen, V.B. (1985) A re-examination of feedback in computer-based instruction: implications for instructional design. Educational Technology, 25(1), 33-37.

Davis, G.A. (1966) Current status of research and theory in human problem solving. Psychological Bulletin, 66(1), 36-54.

Dunker, K. (1945) On problem-solving. Psychological Monographs, 32 (5 Whole No. 270).

Ericsson, K.A. & Simon, H.A. (1984) Protocol Analysis: Verbal Reports as Data. The MIT Press, Cambridge, Mass.

Fetter, W.R. (1984) Guidelines for evaluation of computer software (with an evaluation form). Educational Technology, 24(3), 19-21.

Glaser, *B.G. & Strauss,* A.L. (1967) The Discovery of Grounded Theory: Strategies for Qualitative Research. Aldine, New York.

Hawkins, R.P. & Dotson, V.A. (1975) Reliability scores that delude: an Alice in Wonderland trip through the misleading characteristics of interobserver agreement scores in interval recording. In Behavior *Analysis: Areas of Research and* Application. In (eds Ramp, E. & Semp, G.) pp.359 - 376. Prentice Hall, Englewood Cliffs, NJ.

Jay, T.B. (1983) The cognitive approach to computer courseware design and evaluation. Educational Technology, 23(1), 22-26.

Johnston, V.M. (1987) The evaluation of microcomputer programs: an area of debate. Journal of Computer Assisted Learning, 3(1), 40-50.

Larkin, J., McDermott, J., Simon, D.P. & Simon, H.A. (1980) Expert and novice performance in solving physics problems. Science, 208, 1335 - 1342.

Lewicki, P., Czyzewska, M. & Hoffman, M. (1987) Unconscious acquisition of complex procedural knowledge. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13(4), 523-530.

Maier, N.R.F. (1945) Reasoning in humans: the mechanisms of equivalent stimuli and of reasoning. Journal of Experimental Psychology, 35, 349-360.

Preece, J. & Jones, A. (1985) Training teachers to select educational software: results of a formative evaluation of an Open University Pack. British *Journal* of Education Technology, 16(1), 9-20.

Reid, J.W. (1951) An experimental analysis of 'Analysis of the Goal' in problem solving. *Journal* of *General* Psychology, 44, 51-69.

Roblyer, M.D. (1985) The greening of educational computing: a proposal for a more research-based approach to computers in instruction. Educational Technology,
25(1), 40-44.

Salomon, G. & Gardner, H. (1986) The computer as education: lessons from television research. Educational Researcher, 15(1), 13-19.