STRUCTURAL VARIABLES AFFECTING CAI PERFORMANCE ON ARITHMETIC WORD PROBLEMS OF DISADVANTAGED AND DEAF STUDENTS

by

Barbara W. Searle, Paul Lorton, Jr.,

and Patrick Suppes

TECHNICAL REPORT NO. 213

September 4, 1973

PSYCHOLOGY AND EDUCATION SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

Structural Variables Affecting CAI Performance on Arithmetic Word Problems of Disadvantaged and Deaf Students (1)

Barbara W. Searle, Paul Lorton, Jr.,

and Patrick Suppes

Introduction

A central theme of mathematics instruction is to help students develop problem-solving skills that generalize beyond the tasks of elementary-level mathematics curriculums. Instruction in solving arithmetic word problems is one method of teaching problem-solving skills. Word problems are difficult for most students, and despite intense interest and investigation, much remains to be learned about the sources of problem difficulty. Using the capabilities of the computer, we have designed an instructional program that emphasizes students' problem-solving skills instead of their computational skills, and that allows the collection of a large and detailed data base.

The study reported here has three purposes: (a) to identify structural variables that affect performance of students on arithmetic word problems presented at a computer terminal, (b) to use the identified variables to structure a computer-based problem-solving curriculum, and (c) to assess the usefulness of the identified

(1) This research was supported by Office of Education Grant OEG-0-70-4797 (607) and NSF Basic Research Grant GJ-443X.

variables as predictors of student performance on the newly structured curriculum.

The study was conducted in two phases. During the first phase 700 arithmetic word problems were written and edited. Predicting problem difficulty on the basis of results from pilot studies that used multiple linear-regression models, we structured the problem-solving curriculum by ordering problems from the least difficult to the most difficult. During the second phase, students who were enrolled in a computer-assisted instruction (CAI) arithmetic program given by the Institute for Mathematical Studies in the Social Sciences (IMSSS) of Stanford University were also given problem-solving (PS) lessons. The student population was drawn from several schools for the deaf and from a school in an economically disadvantaged area. Using the performance data collected for these students to reanalyze problem difficulty, we revised the set of structural variables and made new predictions for the problem set, which permitted a restructuring of the curriculum.

Description of the Problem-solving Course

The PS course is designed to give students practice in solving arithmetic word problems. Tutorial functions of the course are limited to giving general and, in some cases, specific hints. The emphasis of the course is on methods of solution; the student constructs a well-formed algebraic expression, but it is the computer that carries out the computations. Thus, the student learns a set of simple commands that tell the computer which computations to carry out. Although the text of the problems is stored by the computer, the numbers used in each problem are generated for each presentation. As a result, students who see the same problem statement have different numbers to work with.

In solving problems, the student is free to experiment with the computer calculator made available to him. The computer evaluates his response only when he instructs it to do so. The student is given three chances to respond correctly to a problem. After the third error, the computer types the correct numerical answer to the problem, and then types the problem again with newly generated numbers. The student may skip a problem at any time, and, as a result, is not obliged to solve every problem. Examples of problem output with student input are shown in Figure 1. Student input is underlined. The student uses the equal sign (=) to request evaluation.

Insert Figure 1 about here

The student constructing a solution in the PS course is also free to use any combination of steps. The computer calculates the correct answer from a stored solution string using the numbers generated for the problem presentation, and compares the result with the student's answer.

```
---------------
PROBLEM I103

TOM HAS 63 CARS.  HE GIVES 4 TO MAX.  HOW MANY CARS DOES
TOM HAVE NOW?
         A =       63
         B =        4


*?

TRY SUBTRACTING.

*A-B=
         C =       59
         WELL DONE

---------------
PROBLEM I106

MR. BROWN HAS $9830 IN ONE BANK AND HE HAS $590 IN ANOTHER BANK,
TOO.  HOW MUCH MONEY DOES HE HAVE IN THE TWO BANKS?
         A =     9830
         B =      590


*AXB=
         C = 5799700
         NOT QUITE, TRY AGAIN.

*A+B=
         D =    10420
         BETTER

---------------
PROBLEM I109

JUAN HAS 99 PET FISH, 3 PET CATS, AND 7 PET RABBITS.
HOW MANY PETS DOES HE HAVE IN ALL?
         A =       99
         B =        3
         C =        7


*A+B
         D =      102

*D+C=
         E =      109
         GROOVY

---------------
PROBLEM I118

DANIEL HAD 91 ROCKS AND HE GAVE THEM TO 5 CHILDREN.
HOW MANY ROCKS DID EACH CHILD GET?
         A =       91
         B =        5


*A/B=
         C =       18.200
         GREAT
```

Fig. 1. Sample output from the PS course.

Pilot Study

Data from studies by Loftus and Suppes (1972) and Suppes, Loftus, and Jerman (1969) were used to calculate a linear regression model for performance on arithmetic word problem solving. These studies were based on 100 word problems of appropriate difficulty for sixth-grade students. Sixty-five of the 100 problems were completed by all subjects in both studies and were used in the present analysis.

Subjects in the Suppes, Loftus, and Jerman study (Group 1) were 27 students from an accelerated mathematics class composed of fifth graders from several upper middle-class elementary schools. Subjects in the Loftus and Suppes study (Group 2) were 16 sixth graders from two schools in a culturally disadvantaged area. The two groups performed quite differently. The mean percentage correct on the set of 65 problems was 85.0 (SD = 17.6) for Group 1 and 56.7 (SD = 28.5) for Group 2. In the present analyses, data for both groups were pooled. The mean percentage correct for the Pooled Group was 74.5 (SD = 20.0).

Variables Characterizing Problem Difficulty

The variables chosen to characterize problem difficulty are listed and defined in Table 1. These variables describe aspects of arithmetic word problems. Because this study emphasizes developing a curriculum, the variables are not exhaustive; instead they encompass major features of problem structure. A more detailed set of variables is presented in Jerman (1971). Most of the variables fall into one of two groups: those that describe a standard solution algorithm for the problem, and those that describe the textual statement of the problem. A few variables depend for their definition on mathematical aspects of the problem that cannot be unambiguously identified in a solution algorithm.

Insert Table 1 about here

A standard solution algorithm was written for each problem. For most problems such an algorithm could be specified unambiguously. Where two or more different, but correct, algorithms could be constructed, the choice depended on (a) the method presented for solving the problem in standard elementary-level mathematics texts, (b) the intuitive judgment of the authors based on their experience with children's problem-solving behavior, and (c) the variable values assigned to the algorithm. When two algorithms were judged 'natural', using the criteria described in (a) and (b), the algorithm that gave a minimal sum of the variable values was chosen.

Variables that describe characteristics of the solution algorithm are OPERS, STEPS, ORDER, ADD, SUB, MUL, and DIV. The variables OPERS, ORDER, and STEPS are not independent; the value of STEPS places an upper limit on the possible values of OPERS and ORDER.

Variables that describe the textual statement of the problem are LENGT and VCLUE. A simple measure of verbal complexity, number of words in the problem statement (LENGT), was chosen for this study.

TABLE 1

Definition of Variables Used for Pilot Study

Variable	Name	Range	Definition
x_1	OPERS	1-3	Minimum number of different arithmetic operations required to reach a solution.
X_2	STEPS	1-7	Minimum number of binary operations required to obtain an answer.
X_3	LENGT	7-51	Number of words in the problem. Each number symbol counts as one word.
X_4	CONVR	0,1	Problem is said to have a conversion (coded 1) if conversion of units is required and the equivalent units are not presented in the problem statement.
X_5	VCLUE	0,1	Problem has a verbal clue (coded 0) if (a) there is a clue for each required operation, and (b) if the clue word (or phrase) is one of the following: for +, added, altogether, gained; for -, how much (less, more and synonyms); for x, each; for /, average.
X_6	ORDER	0,1	Order is the same (coded 0) if the numbers in the problem are presented in the same order as they occur in the coded solution string.
X_7	FORMU	0,1	Solution of the problem requires knowledge of a formula not included in the problem presentation (coded 1).
X_8	AVERG	0,1	The word average is in the problem statement, and the student must compute an average or use an average to solve the problem (coded 1).
X_9	ADD	0,1	Solution requires an addition.
X_{10}	SUB	0,1	Solution requires a subtraction.
X_{11}	MUL	0,1	Solution requires a multiplication.
X_{12}	DIV	0,1	Solution requires a division.

Although Loftus and Suppes (1972) reported the advantages of including a measure that characterizes the structural complexity of sentences, difficulties of coding this measure precluded its use in this study. The variable VCLUE, which indicates the presence of a verbal clue in the problem statement, depends on both verbal and mathematical properties of the problem.

The variables CONVR and FORMU describe problems that require, respectively, conversion of units and knowledge of a formula. Frequently problems of these types call for the student to use a number, a conversion factor, or other constant not presented in the problem statement. The same is true of the variable AVERG, which also requires the presence of the word 'average' in the problem text.

The Regression Model

A stepwise, multiple regression-analysis program (Dixon, 1970), adapted for the Institute's PDP-10 computer system, calculated regression coefficients, standard errors of estimate, multiple correlation coefficients (R), and the square of the multiple correlation coefficients for the 65 problems completed by the Pooled Group. Proportion correct was the dependent variable in these regressions. Suppes, Loftus, and Jerman (1969) describe the regression model in detail.

The regression equation was

$$z_i = -1.80 + .18X_{i1} + .02X_{i2} + .01X_{i3}^{**} + .37X_{i4}^{*} + .34X_{i5}^{**} + .04X_{i6} + .83X_{i7}^{**} + .11X_{i8} + .05X_{i9} - .08X_{i11} + .26X_{i12}^{*},$$

*p < .05;
**p < .005;

with a multiple R of .81, a standard error of estimate of .36, and an R^2 of .66. The order in which variables were entered into the stepwise regression is presented in Table 2. Six variables, OPERS, VCLUE, DIV, LENGT, FORMU, and CONVR, accounted for 64 percent of the dependent variable variance. The variables FORMU and CONVR characterized few of the problems; the mean coding for FORMU was .03 and for CONVR was .08. These two variables were combined into a single variable, CONFO (X_4), which was coded as 1 if the problem solution required either a conversion or knowledge of a formula.

Insert Table 2 about here

The regression equation, using the five variables, OPERS, LENGT, CONFO, VCLUE, and DIV, was

(1) $$z_i = -1.79 + .23X_{i1}^{*} + .02X_{i3}^{*} + .46X_{i4}^{*} + .27X_{i5}^{*} + .34X_{i12}^{*},$$

*p < .005;

with a multiple R of .78, a standard error of estimate of .37, and

TABLE 2

Order of Introduction of the Variables

in the Regression (Pilot Study Data)

Variable		Multiple R	Standard error of estimate
X_1	OPERS	.66	.43
X_5	VCLUE	.70	.41
X_{12}	DIV	.73	.40
X_3	LENGT	.76	.38
X_7	FORMU	.79	.36
X_8	CONVR	.80	.36
X_9	ADD	.81	.36
X_8	AVERG	.81	.36
X_{11}	MUL	.81	.36
X_2	STEPS	.81	.36
X_6	ORDER	.81	.36

an R^2 of .60. Table 3 presents the regression coefficients, T values, and partial correlation coefficients computed for each of the five independent variables.

Insert Table 3 about here

Construction of the Curriculum

Equation 1 was used to predict the probability correct for each of the 700 problems written for the PS course. The probabilities obtained ranged from .95 to .07. Using the calculated probabilities, we constructed the curriculum by ordering the problems from easiest to hardest. In addition to the 700 ordered problems, 39 introductory problems were written to instruct students on interacting with the program. Fourteen nonnumerical problems taught the students to find characters on the teletypewriter keyboard, to ask for a hint, and to request an evaluation of an answer. Twenty-five numerical problems illustrated different problem types and ranged in predicted difficulty level from .79 to .95.

TABLE 3

Regression Coefficients, Standard Errors of Regression Coefficients and Computed T Values (Pilot Study Data)

Variable	Regression coefficient	SE	Computed T value
X_1 OPERS	.233	.077	3.026
X_3 LENGT	.017	.005	3.400
X_4 CONFO	.459	.158	2.905
X_5 VCLUE	.272	.099	2.747
X_{12} DIV	.337	.107	3.149

Subjects

The experimental subjects were fourth, fifth, and sixth graders enrolled in the IMSSS arithmetic CAI course. Approximately two-thirds of the students came from a primarily black California elementary school and the remainder came from schools for the deaf in several parts of the country.

The black students were from an economically depressed area in Santa Clara County, California, where the school district comprises 5 percent of the total county school population. Of the entire population of county welfare families, 35 percent live within the school district. Students in Grades 4-6 are, on the average, from one to three years below grade level in arithmetic computation skills.

The majority of the deaf students were enrolled in residential schools for the deaf in several parts of the country. The degree of hearing loss among the students (at least 60 decibels in the better ear) was that adopted for admission standards by the participating schools. Such deaf students are, on the average, from two to three years below grade level in arithmetic computation skills.

Each student took arithmetic lessons at a teletypewriter terminal connected to the IMSSS PDP-10 computer system by telephone lines. A student became eligible for the PS course when his average grade placement on the CAI arithmetic program reached 4.0. Thereafter, if his teacher chose to enroll him in PS, he received a PS session every fifth day. Thus each student started the course at a different time of year and proceeded at his own pace through the curriculum. Of approximately 300 students who received some portion of the course, 120 completed the introductory problems. Approximately 50 students in this group went on to complete the first 100 ordered problems. The data reported here are for 125 problems, the 25 numerical introductory problems, and the first 100 problems of the ordered set. From 51 to 309 responses were recorded for individual problems.

Results

Although the students in this study came from two very different disadvantaged populations, their performance was similar in this setting. Mean values for six performance measures for deaf and hearing students are presented in Table 4. Two measures are for responses that were correct on the first try: the number of steps used to reach a correct solution, and the time in minutes from the completion of the problem presentation at the terminal to the request for evaluation by the student. The measures for incorrect answers record the time and number of steps used by the student to complete the problem when his first response was incorrect. Recall that the student was given a maximum of three opportunities to have his answer evaluated. Also included are the proportion of correct responses and the proportion of problems for which a hint was requested. The deaf and hearing students did not differ on any of these performance measures. In addition, there is a significant correlation between the rank-order of problems for the two groups (Kendall's rho = .511, $p < .001$), indicating that both groups found the same problems easy or hard.

Insert Table 4 about here

Our finding of similarity between two disadvantaged populations is significant. The students whose responses were examined do not represent random samples of the two disadvantaged groups, since eligibility for the course depended on a minimal performance level, and all the students did not complete the same number of problems. Presumably, those students least able to cope with the course dropped out earliest. Nevertheless, the types of handicaps characterizing the two groups do not seem to produce differential performance in this setting. For all further discussion of experimental results, data for the two groups were pooled.

The proportion correct for each problem was obtained, and the distribution of these proportions is shown in Figure 2. Although predicted probability correct for the 125 problems used in the analysis ranged from .79 to .95, the observed proportions ranged from .03 to .94. Moreover, a comparison of problem order for observed, and predicted proportions correct indicated that the rank of the observed values was random with respect to the previously established ranking (Kendall's rho = -.086). The proportion of correct responses for 70 problems fell in the range .60 to .94. For all but 5 problems, the

TABLE 4

Comparison of Performance Measures

for Deaf and Hearing Students

Measure	Mean	
	Deaf	Hearing
Proportion correct	.692	.706
Number of steps for correct solution	1.170	1.129
Latency for correct solution (min.)	.442	.422
Number of steps for incorrect solution	3.045	3.036
Latency for incorrect solution (min.)	1.221	1.274
Proportion of hints requested	.179	.141

observed proportion correct was lower than predicted. The mean difference between observed and predicted proportions was -.22. Thus, the pilot study overestimated student performance, but this is hardly surprising in view of the superior ability of some of the students whose response data constituted part of the pilot study.

Insert Figure 2 about here

There were 21 problems for which the difference between the predicted and observed proportions was greater than -.40. Ten of these were introductory problems and some poor performance could be accounted for by the unfamilarity of problem types selected for illustration. An examination of the remaining 11 indicated that the range of the variable ORDER should be expanded and that more attention should be given to the length of words in problem statements.

Two particularly difficult problems were, "What number divided by # gives #?" and "What fraction of # is #?" (The # is replaced by a program-generated number in presenting the problem.) One possible explanation for the difficulty of these problems is the terseness of the statement and the absence of a setting or 'story'. This suggested that a new variable be defined to distinguish between 'algebraic' and 'story' problems.

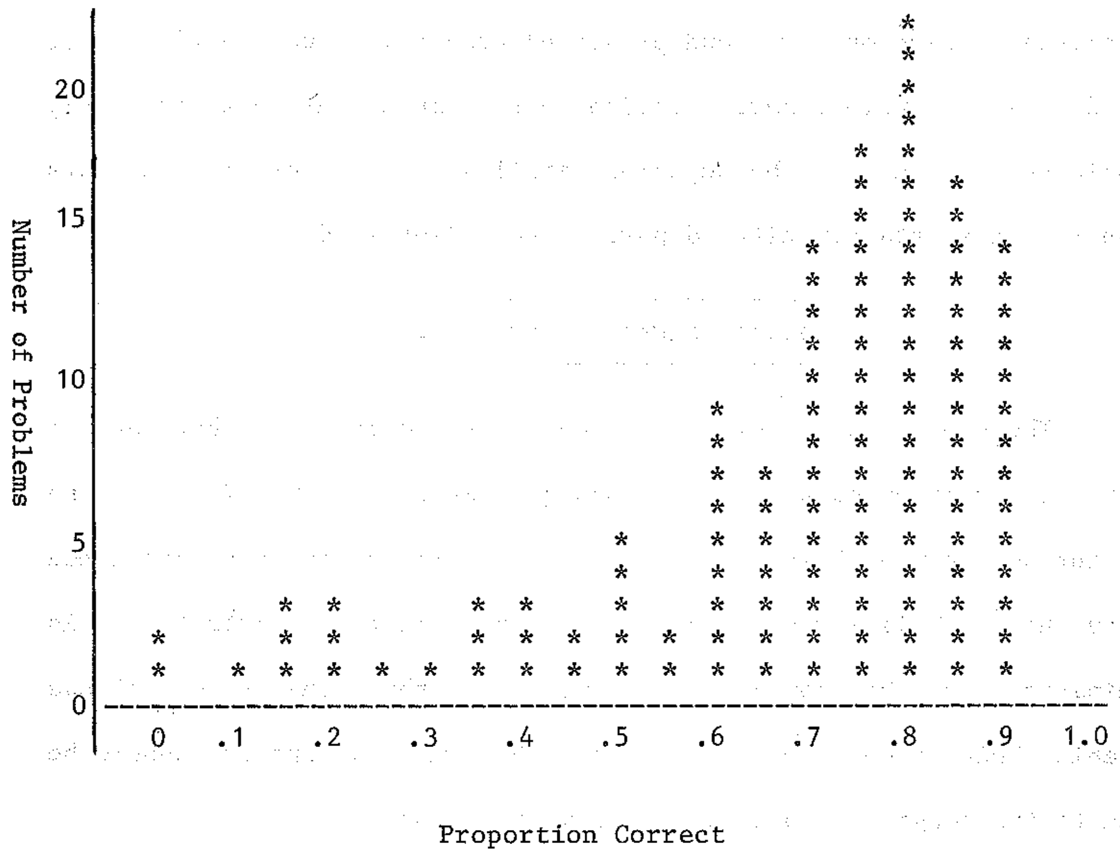

Fig. 2. Frequency distribution of average proportion correct for 125 PS course problems.

A regression analysis for all 125 problems was recalculated, using observed proportion correct as the dependent variable and the same independent variables as in Equation (1). The regression equation was

$$z_i = -1.80 + 1.19X_{i1}^{**} - .001X_{i3} + 1.36X_{i4}^{*} + .21X_{i5}^{*} + .85X_{i12}^{**}$$

$^{*}p < .005$;
$^{**}p < .001$;

with a multiple R of .66, a standard error of estimate of .38, and an R^2 of .44, which was considerably lower than the R^2 of .60 obtained with the pilot study data.

Table 5 presents regression coefficients, T values, and partial correlation coefficients for each of the five independent variables. A comparison of these results with those in Table 3 shows the increased contribution of OPERS, CONFO, and DIV, and the decreased contribution of LENGT to predicted probability correct.

Insert Table 5 about here

The regression coefficient for LENGT was no longer significantly different from zero. This finding is surprising considering the language difficulties of deaf children, who constituted approximately one-third of the student group. Problem length in the pilot study ranged from 7 to 51 words (mean = 29.61, SD = 9.69) and in the PS curriculum from 7 to 39 words (mean = 21.14, SD = 8.73).

TABLE 5

Regression Coefficients, Standard Errors of Regression Coefficients, and Computed T Values (PS Course Data)

Variable	Regression coefficient	SE	Computed T value
X_1 OPERS	1.194	.176	6.767
X_3 LENGT	-.001	.005	.184
X_4 CONFO	1.362	.479	2.841
X_5 VCLUE	.207	.072	2.860
X_{12} DIV	.851	.164	5.195

Thus, problems in the PS curriculum were generally shorter. However, this difference does not seem sufficient to account for the independence of probability correct and problem length. The four remaining variables have significant regression coefficients, and, except for VCLUE, are substantially larger than the corresponding coefficients from the pilot study analysis.

Before continuing the analysis, the definition of several variables was sharpened, others were divided into two variables, and several new variables were defined. The variables used for further analysis of the PS data are shown in Table 6.

Insert Table 6 about here

Thirteen additional variables were defined. Three of these, MAXWD, MAXSN, and NUMSN, characterize the problem statement in greater detail than LENGT. After an examination of the raw data, some specific variables (19-22) describing the structure of subtraction problems were defined. The variables SEQUE and POSIT attempt to account for the position of a problem in relation to neighboring problem types (SEQUE) and for the amount of practice the student is likely to have had (POSIT).

The multiple regression analysis was repeated using 23 variables. There were no problems exemplifying FORMU and AVRG1. This analysis yielded a multiple R of .90, a standard error of estimate of .24, and an R^2 of .81. The order of variables entered in the

TABLE 6

Definition of Variables Used for Analysis of PS Course Data

Variable	Name	Range	Definition
X_1	OPERS*	1-2 1-4	Number of different arithmetic operations required to reach a solution, using the coded solution string.
X_2	STEPS*	1-3 1-9	Number of binary operations required to obtain an answer, using the coded solution string.
X_3	LENGT	7-79	Number of words in the problem. Each number symbol (#) counts as one word.
X_4	CONV1	0,1	Problem is said to have a conversion (coded 1) if conversion of units is required and the equivalent units are not presented in the problem statement.
X_5	VCLUE*	0,1	Problem has a verbal clue (coded 0) if (a) operation is + and problem has word 'together' or 'altogether', or if (b) operation is - and problem has phrase 'have left' or 'were left', or if (c) operation is X and problem has word 'each'.
X_6	ORDER*	0-2 0-3	The number of adjacent pairs of letters in the solution string that are not in alphabetical order.
X_7	FORMU	0,1	Solution of the problem requires knowledge of a formula not included in the problem presentation (coded 1).
X_8	AVRG1*	0,1	The word average is in the problem statement, and the student must compute an average (coded 1).
X_9	ADD	0,1	Solution requires an addition.

Note.--Range printed in brackets characterizes full 700-problem set.
*Definition different from that presented in Table 1.

TABLE 6, cont.

Variable	Name	Range	Definition
X_{10}	SUB	0,1	Solution requires a subtraction.
X_{11}	MUL	0,1	Solution requires a multiplication.
X_{12}	DIV	0,1	Solution requires a division.
X_{13}	MAXWD	5-14 4-16	Length of longest word in problem.
X_{14}	MAXSN	5-30 6-37	Number of words in longest sentence.
X_{15}	NUMSN	1-4	Number of sentences.
X_{16}	CONV2	0,1	Problem requires a conversion of units and equivalent units are presented in the problem (coded 1).
X_{17}	ALGER	0,1	Problem statement is an algebraic statement, not a 'story' (coded 1).
X_{18}	CONST	0-1 0-4	The number of constants in the coded solution string. (Overlaps AVERG and CONV1.)
19-22			Type of subtraction problem (coded 1).
X_{19}	SUBT1	0,1	Type 1: Have a, take away b. How many left?
X_{20}	SUBT2	0,1	Type 2: Have b. How many more do you need to make a?
X_{21}	SUBT3	0,1	Type 3: "b" + "c" = a. "b" = b. Therefore "c"=?
X_{22}	SUBT4	0,1	Type 4: "a" - "c" = b. "a" = a. Therefore "c"=?
X_{23}	POSIT	1-4	Position in problem set, problems # 1-25 coded 4, 26-50 coded 3, 51-75 coded 2, 76-100 coded 1.
X_{24}	SEQUE	0,1	Coded 1 if solution string of preceding problem is exactly same as current problem.
X_{25}	AVRG2	0,1	The word 'average' is in the problem statement and student must use an average to solve the problem.

regression, the multiple R, and the stepwise increase in R^2 are presented in Table 7.

Insert Table 7 about here

Seven variables increased R^2 more than 1 percent. These were ORDER, OPERS, ALGER, ADD, SUBT1, DIV, and STEPS. Three of these variables contributed significantly to the prediction of probability correct for the pilot study. Of the newly defined variables, only ALGER and SUBT1 are included in this group. It is interesting that no variables characterizing word and sentence length contributed significantly to the regressions. Because SUBT1 was highly correlated with SUB (r = .73), SUB was used in place of SUBT1 in further analyses. The regression equation using the variables ORDER, OPERS, ALGER, ADD, SUB, DIV, and STEPS was

$$z_i = -1.76 + 1.10X_{i1}^{**} + .19X_{i2}^{*} + .81X_{i6}^{**} + .35X_{i9}^{**} - .18X_{i10}^{*} + .35X_{i12}^{*} + .43X_{i17}^{*}$$

*p < .005;
**p < .001;

with a multiple R of .85, a standard error of estimate of .27, and an R^2 of .73. Thus, nearly 75 percent of the variability in student response to 125 problems in the PS course was accounted for by seven structural variables.

TABLE 7

Order of Introduction of the Variables

in the Regression (PS Course Data)

Variable	Multiple R	Increase in R^2
X_6 ORDER	.571	.326
X_1 OPERS	.734	.212
X_{17} ALGER	.780	.070
X_9 ADD	.814	.054
X_{19} SUBT1	.846	.052
X_{12} DIV	.857	.018
X_2 STEPS	.867	.018
X_{16} CONV2	.873	.010
X_{23} POSIT	.878	.009
X_5 VCLUE	.885	.012
X_{24} SEQUE	.891	.009
X_{18} CONST	.892	.002
X_{14} MAXSN	.893	.001
X_{13} MAXWD	.894	.001
X_3 LENGT	.895	.001
X_{15} NUMSN	.895	.001
X_{22} SUBT4	.896	.000
X_{20} SUBT2	.896	.000
X_{11} MUL	.896	.000
X_{21} SUBT3	.899	.006
X_4 CONV1	.900	.000
X_{25} AVRG2	.900	.000

Summary

We have shown that it is possible to account for a substantial portion of variability in student responses to arithmetic word problems using variables that describe structural features of the problems. However, the results obtained at this stage in our investigations are situation-dependent. The greater variance in observed proportion correct compared with the variance in the predicted proportions, and the differing sets of variables contributing significantly to the regressions come as no surprise. First, the population for this study differed from that used in the pilot study. Second, it is clear that characteristics of the problem set, for example, the frequency of occurrence of exemplars for the range of values for each variable and the way variable values are combined in problem types, affect the weighting for each variable in the regression analysis. Thus, differences were expected because different problem sets were used for the pilot study and the present study.

In the light of these differences, the similarity in performance of the two disadvantaged groups gains in significance, and deserves further study. We believe we can increase the generalizability of our results by redesigning the basic problem set to exemplify in a balanced fashion the full range of variables found to account for problem difficulty. Given, however, the difficulty of making accurate predictions about problem-solving results, the correctness of this belief needs to be explicitly tested.

References

Dixon, W. J. BMD Biomedical Computer Programs. Los Angeles: University of California Press, 1970.

Jerman, M. Instruction in problem solving and an analysis of structural variables that contribute to problem-solving difficulty. Technical Report No. 180. Stanford: Institute for Mathematical Studies in the Social Sciences, Stanford University, 1971.

Loftus, E., & Suppes, P. Structural variables that determine problem-solving difficulty in computer-assisted instruction. Journal of Educational Psychology, 1972, 63, 531-542.

Suppes, P., Loftus, E., & Jerman, M. Problem solving on a computer-based teletype. Educational Studies in Mathematics, 1969, 2, 1-15.