TRANSLATING BETWEEN SYMBOL SYSTEMS: ISOLATING A COMMON DIFFICULTY IN SOLVING ALGEBRA WORD PROBLEMS*

John Clement
Jack Lochhead
and Elliot Soloway

TR # 79-19

Cognitive Development Project
Department of Physics
and Astronomy
University of Massachusetts
Amherst, Massachusetts

March, 1979

*Research reported in this paper was supported by National Science Foundation grant #SED-77-19226; and by NSF-NIE Joint Research Program Grant #SED 78-22043; and the Army Research Institute for the Behavioral and Social Sciences, ARI #DAH3 19-77-G-1002.

ABSTRACT

Many science oriented college freshmen cannot solve a particularly important kind of algebra word problem. The major source of difficulty is the translation process between words and equations; it is not in the ability to comprehend English or manipulate algebra. Meaningful translations between symbol systems require a more complex process than previously recognized.

Humans and certain other primates can use a variety of different symbol systems such as spoken language, sign language, mime, writing, and mathematics. It is often tacitly assumed that if an individual has mastered syntactical rules within each of these systems he will be able to translate between any two of them, but there are reasons to believe that this is not true. For example, foreign language instruction emphasizes grammar and vocabulary; yet many grammatically correct translations by translators not familiar with the subject matter do not convey the appropriate meaning. Also in machine translation of natural languages, purely syntactic translation algorithms have proved to be inadequate to the task. (1)

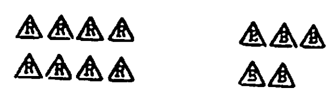
Paige and Simon have shown that many people depend on syntactic strategies when they translate English word problems into algebraic equations, but that while these rules are adequate for some problems they can produce incorrect or meaningless results in others. (2) The data we present confirm these findings and expose a class of problems which should be trivial for a scientifically literate person but which are solved incorrectly by large numbers of science-oriented students.

Table 1 shows selected problems from a 45 minute, written test that was given to 150 freshman engineering students at a major state university. The test was administered during a regularly scheduled class period early in the first semester. Subjects were told that their performance would not affect grades but that the test would help us determine how to improve engineering instruction. All appeared to take the test seriously and all finished their work in the allotted time.

Items 1, 2, and 3 were designed to test algebraic skills. For each problem, over 90% of the students were able to manipulate these algebraic expressions correctly. Items 4, 5 and 6 tested the ability to read written English and translate it into a representation suitable for simple numerical

Table 1

Test Questions (n = 150)

	Question	Correct answer	% correct	Typical wrong answer
1.	Solve for x: $5x = 50$	$x = 10$	99	
2.	Solve for x: $\frac{6}{4} = \frac{30}{x}$	$x = 20$	95	
3.*	Solve for x in terms of a: $9a = 10x$	$x = \frac{9a}{10}$	91	
4.*	There are 8 times as many men as women at a particular school. 50 women go to the school. How many men go to the school?	400	94	
5.	Jones sometimes goes to visit his friend Lubhoft driving 60 miles and using 3 gallons of gas. When he visits his friend Schwartz, he drives 90 miles and used _?_ gallons of gas. (Assume the same driving conditions in both cases.)	4 1/2	93	
6.	At a Red Sox game there are 3 hotdog sellers for every 2 Coke sellers. There are 40 Coke sellers in all. How many hotdog sellers are there at this game?	60	93	
7.	Write an equation using the variables S and P to represent the following statement: "There are six times as many students as professors at this University." Use S for the number of students and P for the number of professors.	$S = 6P$	63	$6S = P$
8.	Write an equation using the variables C and S to represent the following statement: "At Mindy's restaurant, for every four people who order cheesecake, there are five people who ordered strudel." Let C represent the number of cheesecakes and S represent the number of strudels ordered.	$5C=4S$	27	$4C = 5S$
9.**	Write an equation of the form P_A = ____ for the price you should charge adults to ride your ferry boat in order to take in an average of D dollars on each trip. You have the following information: Your customers average 1 child for every 2 adults; Children's tickets are half-price; Your average load is L people (adults and children). Write your equation for P_A in terms of the variables D and L only.	$\frac{6D}{5L}$	2	most gave up
10.*	Write a sentence in English that gives the same information as the following equation : $A = 7S$. A is the number of assemblers in a factory. S is the number of solderers in a factory.		29	***
11.*	Spies fly over the Norum Airplane Manufacturers and return with an aerial photograph of the new planes in the yard.  They are fairly certain that they have photographed a fair sample of one week's production. Write an equation using the letters R and B that describes the relationship between the number of red airplanes and the number of blue planes produced. The equation should allow you to calculate the number of blue planes produced in a month if you know the number of red planes produced in a month.	$5R = 8B$	32	$8R = 5B$

* n = 34 for these problems ** n=83 for this problem

*** Seven solderers for every assembler

calculations; in each case over 90% were successful. Items 7, 8 and 9 tested the ability to perform increasingly complicated translations from English statements into algebraic statements; 98% failed the most difficult problem (#9) while 37% failed the easiest example (#7). The startling drop in performance from 90% to 60% and below suggest that the students' difficulty can be attributed specifically to the translation process.

The errors made on problems 7 and 8 were largely of one kind; in both cases 68% of the errors were reversals: 6S = P instead of S = 6P and 4C = 5S instead of 5C = 4S. These reversals might be interpreted as careless errors, except that roughly half of the subjects were given the following hint with both problems: "Be careful: some students put a number in the wrong place in the equation." This hint had no significant effect; it increased the percentage of correct solutions by only 3% and 5%, respectively.

To investigate the source of these reversal errors we conducted audio and video-taped clinical interviews with fifteen students who were asked to think out loud as they worked these and other similar problems. In the "Students and Professors" problem, two basic sources of reversals were identified: a syntactic type and a semantic type. In the first, the student simply assumes that the order or contiguity of key words will map directly into the order of symbols appearing in the equation. For example, one student wrote 6S = P and explained, "Well, the problem states it right off: '6 times students'. So it will be six times S is equal to professors."

In the second or semantic type of error, the subject links the equation to the meaning of the problem. However, the equation is seen not as an expression of equivalence but as a description of relative size. To students using this approach, the fact that the "S" side of the equation has a 6 on it indicates that it is larger than the "P" side which has no modifier. Thus, there appear to be more S's than there are P's. For example, one subject wrote

6S = 1P and explained "There's six times as many students, which means it's six students to one professor and this (points to 6S) is six times as many students as there are professors (points to 1P)." When asked to draw a picture to illustrate his equation, the student drew from right to left, one circle with a 'P' in it, an equal sign, and six circles with S's in them. (3) Such subjects interpret the incorrect equation as stating that a large group of students are associated with a small group of professors. In this interpretation the letter 'P' apparently stands for "a professor" rather than "the number of professors" and the equal sign expresses a comparison or association, rather than an equivalence. Thus, although these subjects have an accurate semantic conception of the practical situation, they still fail to symbolize that understanding with the correct equation (see figure 1).

In some protocols, subjects wrote down the correct equation, but then switched to the reversed form. This indicates that for these students the reversed equation is the more compelling one.

In a follow-up study questions 10 and 11 were given to a separate group of 34 students from the same population. About seventy percent of the students produced incorrect answers when translating from an equation to words or from a picture to an equation and over 75% of the errors were reversals. In problem 11 it is difficult to attribute these errors to simply a syntactic strategy; the semantic reversal described above is a more plausible explanation.

It is important to stress that these students have no difficulty in reading English. They are skilled in the manipulation of simple algebraic equations, but when asked to invent a simple equation for a situation they can experience some difficulty. What they cannot do is translate between the two symbolic systems. Most can translate from simple, verbal statements to an equation in one variable, such as (for problem 5):

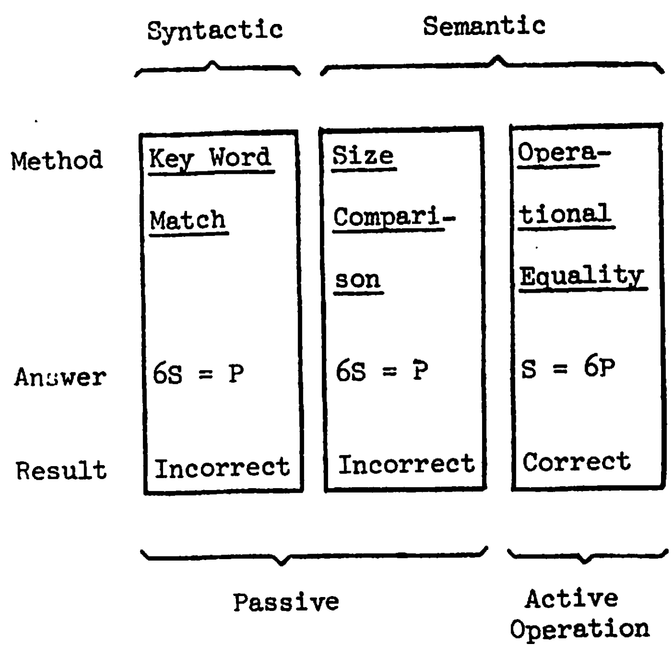


Fig. 1. Solution Methods for "Students and Professors" Problem

$$\frac{60}{3} = \frac{90}{x}$$

but many have difficulty with very simple expressions in two variables.

The structure of the correct translation process is exposed when we clarify certain tacit assumptions underlying conventions in algebraic notation. The correct equation, S = 6P, does not describe sizes of the groups in a literal or direct manner; it describes an equivalence relation that would occur <u>if one were to perform a particular hypothetical operation</u>, namely making the group of professors six times larger than it really is. Some students find the correct equation by writing the reversed equation first and then plugging in numbers as a check. However, analysis of protocols from successful solutions indicates that the key to understanding the correct semantic translation lies in viewing the number six as an <u>operator</u> which transforms the <u>number</u> of professors into the <u>number</u> of students. One subject who correctly wrote S = 6P said, "If you want to even out the number of students to the number of professors, you'd have to have six times as many professors." The equation is thus read as an instruction to act rather than as a static comparison. In this regard we note that because questions 4, 5 and 6 request a numerical result the subject will, at a minimum, carry out an action in the form of an arithmetic operation. This contrasts with questions 7, 8 and 9, where the operations must be carried out implicitly.

In order to investigate the effect of active and static perspectives we examined a question similar to 8 in the context of computer programming. One might expect that writing a computer program is more complicated and hence more difficult than writing an algebraic equation. However, programming, unlike algebra, induces one to take an active procedural perspective. The programmer should: (1) represent all operations explicitly, (2) view the equal sign (=) as an assignment operator, (3) view an equation as a transformation from an input to an output. We felt this perspective might prevent errors of the on translating between symbol systems as a separate skill and in part from the static perspective into which much of mathematics is cast. (4) These results provide a disturbing picture of the level of mathematical understanding commonly attained in technical fields, and they suggest that we need to reevaluate some basic assumptions in mathematics instruction.

John J. Clement

Jack Lochhead

Department of Physics and Astronomy

University of Massachusetts

Amherst, Massachusetts 01003

Elliot Soloway

Department of Computer Science

University of Massachusetts

form described earlier.

Our subjects, in this experiment, were 17 professional engineers, with 10 to 30 years experience. They were taking a one week course on the BASIC programming language. During the first day of the course they were asked to write an equation for the statement: "At the last football game, for every four people who bought sandwiches, there were five who bought hamburgers." Eight of the engineers failed to correctly solve this problem. On the second day of the course, and without any discussion of the answers to the above question they were asked to write a computer program as follows: "At the last company cocktail party, for every 6 people who drank hard liquor, there were 11 people who drank beer. Write a program in BASIC which will output the number of beer drinkers when supplied with the number of hard liquor drinkers." All subjects answered this question correctly using the statement LET B = (11*H)/6 (or some variant) in their program. The form of this statement is equivalent to that of the correct answer to the first question. The success of the engineers in this computational setting supports our earlier hypothesis that the reversal difficulty is associated with viewing the problem from a static perspective.

Fluent translations between symbol systems such as verbal statements, graphs, programs, diagrams and equations are an essential part of scientific thinking. Investigations of the cognitive processes responsible for these translations are still in an embryonic stage. It is well known that many people cannot solve "word problems." We have identified some specific causes of translation errors that locate an important source of this problem. Students who understand the translations discussed in this paper tend to view equations from an active perspective; that is, they see them as describing the result of one or more <u>operations</u>. We believe that the reason so few students reach this level of understanding stems in part from the lack of emphasis schools place

Notes

(1) Among the many articles relevant to this question are: Bobrow, D.G. and Winograd, T., "An Overview of K.R.L., a Knowledge Representation Language," Cognitive Science, 1, 1 (Jan. 1977), and Novak, G.S. Jr., "Representations of Knowledge in a Program for Solving Physics Problems," International Joint Conference on Artificial Intelligence, 286 (1977).

(2) Paige, J. and Simon, H., "Cognitive Processes in Solving Algebra Word Problems," in Problem Solving Research, Method and Theory, B. Kleinmuntz, Ed. (John Wiley & Sons, N.Y., 1966).

(3) Another student, working from the statement: "There are 8 times as many people in China as there are in England," wrote 8C=1E and said, "It means that there is a larger number of Chinese (points to '8C') for every Englishman (points to '1E')."

(4) Detailed analyses of this problem are given by J. Kaput in "Mathematics and Learning" and A. diSessa in "Learnable Representations of Knowledge." Both papers appear in Cognitive Process Instruction, J. Lochhead and J. Clement, Ed. (Franklin Institute Press, Philadelphia, 1979).

(5) We thank R. Narode and N. Fredette for their help in collecting data and F.W. Byron, Jr., S. Polatsek, A. Well and C. Clifton for comments on earlier drafts of this paper.

(6) This work was supported by grants from the National Science Foundation, NSF-RISE #SED77-19226; the NSF-NIE joint research program, #SED78-22043; and the Army Research Institute for the Behavioral and Social Sciences, ARI #DAH3 19-77-G-1002.