Focus on Learning Problems in Mathematics
October 1981, Volume 3, Number 4,

Intuitive misconceptions In Algebra As A Source Of Math Anxiety

John Clement
Ronald Narode
Peter Rosnick

While students sometimes experience difficulties with the rules for simplifying or solving algebraic equations, there is a second category of difficulties in algebraic equations. It is to this second issue that this paper is directed. The rule which allows one to divide both sides of the equation $6P = S$ by six is an example of an equation manipulation rule. On the other hand, the formulation of this equation from a word problem or a data table is a process which requires one to assign a symbolic meaning to the equation.

We are finding that *errors that result from mistranslations of meaning in formulating or interpreting equations* occur frequently even among science-oriented college students. Many of these students can manipulate algebraic equations but cannot reliably formulate or interpret them. Certainly, algebra would be impossible without the logical rules of equation manipulation which characterize the field. However, the value of algebra to the individual must be doubted if it does not convey meaning.

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

Research reported in this paper was supported by NSF Award No. SED 78-22043 in the Joint National Institute of Education/National Science Foundation Program of Research on Cognitive Processes and the Structure of Knowledge in Science and Mathematics.

SUBJECTS

In this paper we describe results from clinical interviews where students are asked to translate from words to equations and from data tables to equations. These interviews have led us to discover a number of *common misconceptions concerning the meaning of variables and equations*. Many of the students interviewed were engineering majors who had either taken calculus or were taking it at the time of the study. Most of these students demonstrated mathematical skills necessary to correctly differentiate complicated functions in the calculus. But in spite of these skills they nevertheless experienced much difficulty with certain simple algebra word problems. We have interviewed approximately fifty students in our studies of algebra word problems. The examples given here are drawn from an exploratory study with 15 students in which we aimed to cataloque a wide variety of solution approaches on different problems. Our subjects were paid volunteers predominantly from freshman physics and calculus courses. All sessions were audio taped and most were videotaped.

OBSERVED BEHAVIOR PATTERNS

Many of the misconceptions identified in this study represent major stumbling blocks, even to the better students. The misconceptions appear not to be simply random errors, but derived from stable conceptual schemata. While some of the ideas are unconventional, they often have convincing "logic" for the student. The behavior patterns discussed in this paper are:

1. Reversal Errors
2. The Lots Approach
3. Focusing on a "Total" by Introducing an Extraneous Variable
4. Vertical Rule Symbolization in Data Table

REVERSAL ERRORS

One of the most prevalent equation formulation errors that we have investigated is a *reversal error* (see Clement, 1980, Clement, Lochhead, and Monk, in press). We will focus most of our attention in this paper on other types of misconceptions but begin with a brief discussion of the reversal error. The following problem, entitled "Students and Professors," is typical of problems that can evoke a reversed equation as a solution.

> 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.

The correct answer to problem is S = 6P. The typical reversal error is 6S = P. One possible cause for this error is that we call "word and order matching" where the student will write an equation in a way that parallels the order of the words in the English sentence. For example, one student described his solution as follows:

S: I thought to myself, um, there's six times as many students equaling the professors. I substituted "as" for equals so I have six students equals professors (referring to his written equation, 6S = P).

It is tempting to suggest that this student's solution illustrates a purely mechanical problem solving process in which he did not "think" about the meaning of the problem text. However, many other students demonstrate their understanding of the relative sizes of the groups of students and professors by drawing pictures, by rewording the problem statement, or by constructing data tables and graphs: yet they frequently choose the equation 6S = P over S = 6P. In these cases their thinking appears to go beyond a simple word order matching approach. It is not just a careless mistake: their choice is one for which they feel they have logical justification.

The following segment of a taped interview is an example of one such justification:

S: I know that there are a lot more, there are more students than there are professors and so the number — this — if I had six by the S it would make it bigger and since there are more students it must be a bigger number.

Here the student understands there are six times as many students but puts this 6 next to the S to symbolize their relationship. Another meaningful way in which the equation 6S = P has been justified is that some students see the letter S as a label standing for 'students' rather than as a variable standing for 'the number of students.' As one student said, "P is a unit (which stands for) one professor."

The reversal error has shown up strongly in other translations as well, not only from words to equations, but from data tables to equations, pictures to equations, and equations to words. (Clement, Lochhead, and Monk, in press). This indicates that the difficulty is not simply a result of the specific wording of the problem text. It appears to be an example of a common and persistent misconception. Furthermore, this mosconception is not easily rectified in tutoring sessions of moderate length (Rosnick and Clement, in press). The reversal error was our first clue to the idea that students might harbor a number of different misconceptions concerning the meaning of algebraic equations.

We will now turn to the major topic of this paper: the observation of a number of new equation formulation errors in addition to reversals. Some of these often occur in a problem in conjunction with a

reversal error, but are actually observable as independent behavior patterns. These patterns are described in the next three sections.

THE "LOTS" APPROACH

A second behavior pattern to be observed in equation formulation tasks occurs when students attempt to solve certain word problems involving ratios in terms of "lots." We have found that a student can use interesting and quasiconsistent methods of reasoning even though they are formulating an incorrect equation. This was apparent in student responses to the following problem:

Write an equation using the variables C and S to represent the following statements: "At Mindy's restaurant, for every four people who order cheesecake, there are five people who order strudel. Let C represent the number of cheesecakes and S represent the number of strudels.

When this equation was administered to 150 university engineering majors, only 27% responded correctly (Clement, Lochhead, Soloway, 1979), and two thirds of the errors were reversals. While it is common to see the reversed equation 4C = 5S it is surpnsing that some students can use this incorrect expression to correctly find one numerical quantity when given the other. Their approach appears to involve thinking of the quantities by corresponding sets of "lots," where groups of four cheesecakes are placed in one-to-one correspondence with groups of five strudels.

This method is illustrated in the following protocol with a freshman engineering major who has written 4C = 5S and is using the incorrect equation to find the correct number of strudels ordered when 20 cheesecakes have been ordered.

S: (Writes 4C = 5S) Well, if 20 people came in and bought cheesecakes — and for every 4 cheesecakes bought there were 5 strudels bought — so — in 20 cheesecakes — you said, right, 20 people bought cheesecakes — so there are 4 sets of 5 and 5 sets of 4 in 20 —

The student now draws 5 groups of x's and 4 x's in each group and the number 5 under each group, as seen in Fig. 1a.

XXXX	XXXX	XXXX	XXXX	XXXX
5	5	5	5	5

Fig. la.

The student continues:

S: And so for every 4 cheesecakes there were 5 strudels bought — The student now draws in arrows from each group of x's to the number 5 saying "this causes that" while drawing each arrow (Fig. 1b.)

XXXX	XXXX	XXXX	XXXX	XXXX
↓	↓	↓	↓	↓
5	5	5	5	5

He then concludes:

S: If you add all these together — you have 25 strudels.

In this protocol, cheesecakes occur in "lots" of four and strudels occur in lots of five. There is a one-to-one correspondence between the two sets of lots. It is interesting that this student introduces a kind of "pseudo-causality," i.e., one lot of four cheesecakes "causes" one lot of five strudels. Equality exists between the number of lots only. It is clear that this student comprehends the numerical relationship described in the problem. It is also clear that this student has misinterpreted the meaning of the variable C, the number of cheesecakes ordered, and S, the number of strudels ordered. His numbers are correct, but his algebraic equation is not.

We have observed this tendency to think in terms of an equal number of groups or "lots" in several students. This is certainly a valid way to think about the problem. However, the students then make an overly direct attempt to symbolize this equality with the incorrect equation $4C = 5S$. Another student doing the same problem said, "I guess C is like how many groups of 4 there are and S is, would be how many groups of 5." Here the letters C and S have lost their specified meaning as variables standing for the number of cheesecakes and the number of strudels.

A third student who worked this problem in a similar fashion was questioned about his use of these variables, at which point he changed his equation from $4C = 5S$ to $(4 * X = C) = (5 * X = S)$ [where X = number of groups]. His explanation follows:

> S: This equation..well, I've got C should equal the number of cheesecakes not the number of groups. So I had to figure another thing to represent the number of groups and that would be X. So...the number of people in the group ..times the number of groups should equal the total number of cheesecakes. The same with over here (points to right side of equation). The number of people in the group — 5 for the 5 strudel, times the total number of groups — times the number of groups, which would be the same as the number of groups here — would equal the total number of strudel. And it should have a 4 to 5 relationship as it does here.

Although the student's equation (with three equal signs!) is incorrect, he has developed a method which will correctly find one quantity when given the other. By introducing a new quantity - "the number of groups" - the four-to-five relationship is exhibited, and for the student the expression "works." This student has the beginning of a correct solution method based on his symbolization of the "equal lots" idea, but he doesn't carry it through to an equation in S and C.

A correct argument along these lines would be as follows:

$$5 \cdot X = S \qquad 4 \cdot X = C$$

$$X = \frac{S}{5} \qquad X = \frac{C}{4}$$

$$\frac{S}{5} = \frac{C}{4}$$

$$4S = 5C$$

Students who generate a reversed equation using the "lots" approach appear to symbolize a correct conceptualization of the problem with an incorrect equation which is consistent with their ideas and may "work" for them, i.e., they can use their equation to output correct values. Such solutions have strong belief-value which may not be eliminated by simply demonstrating the "correct" method. Unless shown *why* their solution is wrong the student may never really learn a correct method. Instead, the student may face the paradox of thinking he understands something which he has been told he doesn't understand. This theme will recur throughout this paper.

FOCUSING ON A "TOTAL" BY INTRODUCING AN EXTRANEOUS VARIABLE

Another phenomenon which has occurred repeatedly is the spontaneous inclusion of a new quantity -- the total. When asked to symbolize with an equation the relationship in the Students and Professors problem some students reply with one of these equations:

(1) $6S + P = T$	(2) $6P + S = T$	(3) $6P + P = T$	(4) $S + P = T$
(wrong)	(wrong)	(right, but not requested)	(right, but insufficient)

The following is a section of a protocol from a student who has been asked to explain his equation, $6S = P$.

> I: What made you think that -- when you wrote that down?
>
> S: It said use the variables "S" and "P." And there are six times as many students as professors. (Hesitation). From what I remember of algebra it should be 6x plus P equal to some number — but I don't have a number here.

It appears, for this student, that the relationship between the number of professors and the number of students only make sense if some kind of *sum* is included in the equation

We have considerd two types of ratio problems: an integral ratio problem where the ratio coefficient may be expressed as an integer (e.g., $6P = S$) and a fractional ratio problem where the coefficient must be expressed as a fraction (e.g., 5/4D = S). Students may also symbolize the relationship between two quantities in a fractional ratio problem by considering their sum. The statement: "At Mindy's restaurant, for every four people who order cheesecake, five order strudel" is sometimes symbolized incorrectly with 4C + 5S = T. There is indeed an implied total in these problems - namely, the number of people "at the university" and the number of pastries "at Mindy's restaurant." However, these "totals equations" were not requested, and most are mistaken (C + S = T and S + P = T are correct). This is an instance where a problem has triggered the activity of more cognitive structures than students need to solve it. It is encouraging that they understand the problem situation well enough to see relationships that are not explicitly stated. Furthermore, some have evidenced an ability to form a synthesis of these separate relationships into a unifying equation. Unfortunately, the equation does not answer the problem of describing the direct ratio of students to professors or cheesecake to strudel. And the totals equations themselves are often wrong in that the students' synthesis of relationships doesn't work, (e.g., 4C + 5S = T instead of C + S = T).

We believe that the spontaneous act of defining the total as a new variable in these problems is a natural consequence of the way the students' ideas are structured. These observations provide important clues concerning the nature of the students' intuitive conceptions for thinking about mathematical situations. Mapping out these conceptions is an important task for future research.

VERTICAL RULE SYMBOLIZATION IN DATA TABLES

Undergraduates also experience difficulty in writing equations from data *tables*. We gave the following problem to 150 freshman engineering students:

> Weights are hung on the end of a spring and the stretch of the spring is measured. The data are shown in the table below. Write an equation using the letters S and W to summarize the data below.

Stretch S (cm)	Weight W (g)
3	100
6	200
9	300
12	400

Fifty-one percent of the students missed this problem. This shows that equations formulation errors are not restricted to translations between words and equations. Twenty of the students gave now answer. Eighteen (32%) of the remaining 57 errors were of the form 3S = 100W or S + 3 = W + 100.

The following analysis of an interview with an upperclass science student illustrates this latter difficulty. The student was given the problem above and immediately attended to the progressions on either side of the vertical line. He began his analysis by examining the *changes* in weight and summarized his findings with $W_1 + 100 = W_2$. Similarly, he described the changes in stretch with $S_1 + 3 = S_2$.

> S: I'm looking at the right side — trying to express the changes there — So it would be the original number plus one hundred on each change. It would be weight one plus 100 will equal weight two. (Writes $W_1 + 100 = W_2$)
>
> I: What is W_1 and W_2?
>
> S: W_1 would be the weight previous to the one you're trying to determine. Now, I want to decide what happened to this side here (points to "stretch" side of table). So it looks like this stretch here — stretch one, to stretch one is added a unit 3 to get stretch two — the next value in the progression. (Writes: $S_1 + 3 = S_2$)

The protocol indicates that these equations are designed to predict the next item in a sequence where at least one item is known. Upon noticing that the problem requested only one equation for the summary of the data table, the student wrote: $S_2 = W_2$, and then substituted his earlier expressions to get $S_1 + 3 = W_1 + 100$. He stated, "this will predict the next entry in the table." His sequence of expressions then looked like the following:

$$W_1 + 100 = W_2$$
$$S_1 + 3 = S_2$$
$$S_2 = W_2$$
$$S_1 + 3 = W_1 + 100$$

The subject's first two equations, apparently, were an attempt to symbolize a vertical rule which predicts subsequent values down the columns of the table. The last equation is a conjunction of the previous three. This intuitive attempt at symbolization is geared toward prediction of subsequent values which necessarily must follow the previous value, i.e., it cannot predict the stretch for an intermediate weight (150 g.). This is in contrast to what might be termed a "horizontal rule" where the functional relationship between variables is sought in an analysis of each row of data. This latter method would yield the equation S = 3/100 W, if done correctly. It should be noted that the student quoted above never indicates a realization that his result is an equation without equality. In fact, it appears that it was never intended to show equality. Rather, the equal sign is used more like an arrow to signify that a value of one variable maps into a different value of the second variable. Its meaning here might best be paraphrased as "corresponds to."

The following transcript excerpt shows how other students will use the same approach of focusing on the "vertical rule" in the data table, but symbolize this with the reversed equation 3S = 100W.

S: It seems that every time you increase this, every time you increase by 100 grams, it stretches three more centimeters. (Writes 3S = 100W).

I: Okay, what are you thinking when you write down 100 there?

S: For every 100 grams, it stretches 3 cm ... Right now it's in direct proportion because..the weight is not sufficient to overstretch it.

I: What would a weight of 600 grams stretch it, if you think it's still within that?

S: 18 cm. You plug in 6 over here [for S in the equation 3S = 100W] and you come out with 18.

I: And ... 1,000 grams?

S: 1,000 grams? 30 cm.

Here the student writes a reversed equation, but is able to generate correct values from it. This is reminiscent of the "lots" approach described earlier. The difficulty this student had in deriving the correct equation stems from the fact that data cannot "speak for itself." The person reading the data must reorganize it into a meaningful pattern. Unfortunately for many students, the particular patterns used to symbolize the data are not obvious and are rarely stated. It is not unusual for a student to symbolize a pattern incorrectly with an equation which has meaning for him. In the case above the equation can actually be used to predict additional data points. Such a student may have difficulty understanding why he is mistaken when he is told that he is mistaken.

The discovery of vertical rule errors and reversal errors in data table problems is important for two reasons. First, data tables play an important role in all aspects of science, especially in laboratory work. Therefore, these errors are a potentially significant source of difficulty in science courses. Second, these results show that semantic errors are not restricted to translations from words to equations. The errors in word problems cannot simply be explained away by saying that they are specific effects of reading comprehension or of wording the problem in a confusing way. The errors show up even when verbal statements are not used.

CONCLUSION

In conclusion, we would like to emphasize two important observations from our research:

1) A student who makes errors on these problems often evidences a clear conception of the problem situation sufficient to produce numerical answers.
2) The student may then use algebraic symbols to symbolize the relationship in his conception of the problem, but he does so in an overly associative manner. His intuitive attempts at symbolization come into direct conflict with the standard interpretation of variables and equations.

These patterns occur in the reversal error as well as in each of the three new misconceptions discussed: symbolizing in terms of "lots," focusing on a "total" by introducing an extraneous variable, and symbolizing the "vertical rule" in a data table. A student who evidences these characteristics may not respond to teaching via admonishment and demonstration. He may have a stable conceptual scheme which he believes in. The teacher may convince the student that he is wrong without helping him to see why he is wrong. Unless a student discovers why his approach does not work he is unlikely to surrender it cognitively. Believing that he understands something and being told otherwise by an instructor can place the student in a disequilibrating or paradoxical situation and can frequently leave him there. Left unresolved, this paradox can produce anxiety — specifically, math anxiety. The student may begin to doubt his own mathematical reasoning skills in general.

Avoiding this all too frequent outcome is perhaps the most compelling motive for paying attention to the intuitive ideas expressed by students. Misconceptions which deal with meaning and understanding like the ones described in this paper, are easy to overlook. Too often, we suspect, the curriculum focuses only on the rules of symbol manipulation. It appears that more attention to equation formulation skills is called for to complement the existing emphasis on equation manipulation skills.

REFERENCES

Clement, J., "Algebra Word Problem Solutions: Analysis of a Common Misconception," Cognitive Development Project, University of Massachusetts at Amherst, 1980. Paper given at AERA conference in Boston.

Clement, J., Lochhead, J., and Monk, G., "Translation Difficulties in Learning Mathematics," in press, *American Mathematical Monthly* (1981).

Clement, J., Lochhead, J., and Soloway, E., "Cognitive Processes Involved in Algebraic Symbolization," technical report, Cognitive Development Project, University of Massachusetts at Amherst, 1979.

Rosnick, P., and Clement, J., "Learning Without Understanding: the Effect of Tutoring Strategies on Algebra Misconceptions, in press, *The Journal of Mathematical Behavior* (1981).

Focus on Learning Problems in Mathematics
October 1981, Volume 3, Number 4.