What Are These Things Called Variables?

By SIGRID WAGNER, University of Georgia, Athens, GA 30602

We all know some of the problems that beginning algebra students have in working with literal symbols, or variables. Two articles in this journal (Herscovics and Kieran 1980; Rosnick 1981) have discussed certain difficulties related to variables and have offered suggestions for overcoming them.

One of my favorate algebra stories highlights another misconception about variables:

> The topic of the lesson was "consecutive integer" word problems. The teacher was trying to prepare students for the x, $x + 1$, ... literal symbolism by starting with a numerical example. "What is the next consecutive integer after 17?" she asked. A student replied, "Eighteen." Knowing that the representation for adding 1 would trouble the students, the teacher sagely asked, "What do we have to *do* to 17 to get 18?" "Add 1," came the reply. "Good," encouraged the teacher. "Now suppose we use x to represent an unknown integer. How can we write the next consecutive integer after x? That is, how can we represent the number we get when we add 1 to x?" Without hesitation, the response was, "y."

This confusion between the linear order of the alphabet and the linear ordering of whole numbers is a common mistake for students who are just being introduced to literal symbols (Wagner 1981). Yet, as this true story succinctly illustrates, even students who are fairly well acquainted with literal symbols can succumb to a seemingly natural temptation when bewildered by a new situation.

The question is, why are literal symbols easy enough for elementary students to use, yet hard enough to cause problems for advanced students? Recent research has identified several factors that make literal symbols easy to use but hard to understand. Many of these factors can be grouped under two main headings:

1. Literal symbols are like numerals, only they are different.
2. Literal symbols are like words, only they are different.

Literal symbols (are) easy to use but hard to understand.

That is, literal symbols have certain characteristics similar to numerals, other characteristics similar to words, and still other characteristics that are uniquely their own.

When students first encounter literal symbols in the elementary grades, it is surely good pedagogy to point out some of the characteristics that these new symbols share with numerals and words—two symbol systems the students are already familiar with. However, some students may then overgeneralize and assume that literal symbols are just a new notation for some old ideas. As time goes by, these students work with literal symbols in increasingly varied and more powerful ways, but they may fail to apprehend the unique properties of literal symbols that give mathematical language much of its precision, generality, and flexibility.

Let us, then, consider some important characteristics of literal symbols. The more characteristics we can identify, the better able we are to devise teaching strategies to help students understand and appreciate these things we call variables.

Letters Are Like Numerals, Only Different

Similarities. Besides the coincidental ordering of letters and numbers already mentioned, there are several other similarities between literal symbols and numerals. For one thing, a few letters, like π and e, actually *are* numerals because they are standard symbols for numbers that have no simple digital representations.

Another similarity between letters and numerals is that they often appear together in mathematical statements. In fact, one of the first ways that students see literal symbols used is in open sentences, like $n + 3 = 17$ or $46 - x = 28$, in which the letters appear right along with numerals, operation signs, and relational symbols. In other words, the letters *look* as though they must behave just like numerals.

Solving these open sentences for "the (true) value" of n or x may suggest to beginning algebra students the most telling similarity between letters and numerals, namely, that a letter can actually serve as a *temporary* numeral, the symbol you write until you figure out what the missing number is and can write the *real* numeral. It is only later, when students learn about replacement sets and truth sets, that the limitations of interpreting letters as numerals become obvious. It is then that students should begin to realize that literal symbols are more than just fancy numerals

Differences. Mathematically speaking, the most significant difference between letters and numerals is the one just alluded to—that numerals represent a single number but letters can represent, simultaneously yet individually, many different numbers, as in $0 < n < 20$ or $y = 3x + 2$. It is this property of simultaneous representation that we refer to when we call literal symbols *variables*, knowing full well that some of these symbols may, depending on the context, represent single unknown numbers or even constants! It is this property of simultaneous representation, in combination with some other characteristics of literal symbols to be mentioned in the next section, that gives mathematical language its capacity for making very general statements—definitions, axioms, theorems, formulas, and so on—in concise and unambiguous form.

Another difference between letters and numerals actually starts out as a similarity, that is, both letters and numerals can be used to identify, or name, many things besides numbers. For instance, we may name a point P or call a function g in much the same way that we identify a post office box by its number. But, there is a subtle difference: generally speaking, identification numbers (numerals) are used to label specific, fixed elements of a set, whereas literal symbol names are more often used to identify random, variable elements. The reason we tend to use letters rather than numerals to represent generalized elements is probably twofold: (1) letters look like abbreviated names and thus are easy even for young students to interpret naively, and (2) we know that letters will be used to represent numerical variables later on, so we do not hesitate to use letters to name other kinds of variable elements as soon as it is convenient and reasonable to do so.

> We could never compile a dictionary of the "meanings" of literal symbols.

Marilyn Matz (1979) has pointed out two other differences between letters and numerals. One is the juxtaposition convention that we use with letters and numerals to indicate multiplication, as in $3mn$, in contrast to the place value interpretation that we give to numerals alone, as in 347. Of course, the reason that the juxtaposition convention works is that letters and numerals come from different symbol systems, and letters can represent numbers having either single-digit or multidigit numerals. Therefore, an expression like $3mn$ can be interpreted unambiguously as the product of 3, m, and n.

The other difference mentioned by Matz is that the signs attached to literal symbols do not always match their value, as they would with numerals. That is, x may represent a negative number and $-x$ may be positive. This sign perversity of literal symbols makes one of the standard definitions for absolute value

$$|x| = x \text{ if } x \geq 0$$
$$|x| = -x \text{ if } x < 0$$

virtually incomprehensible to many students.

Letters Are Like Words, Only Different

Similarities. One outstanding similarity between letters and words is that both can act as placeholders in certain expressions. For example, in the sentence, "*He* is a mathematics teacher," the pronoun *he* can be replaced by the names of different men to obtain verbal statements that are either true or false, just as the x in $x^2 + 2x = 3$ can be replaced by numerals for different numbers to obtain mathematical statements that are either true or false.

This placeholder property is so important that it has led more than one writer to claim that literal symbols are essentially just like verbal expressions (Beberman 1964; Hockett 1967; Russell 1940; Tarski 1941). As we shall see, this claim overstates the similarity between letters and words and understates the power of literal symbols. Nevertheless, students need to understand the placeholder property in order to appreciate the *generality* and *flexibility* of literal symbols, two ideas that we shall consider in more detail shortly.

Another similarity between letters and words has already been mentioned, namely, that letters are often chosen to suggest abbreviations for words as, for example, when n is used to represent the missing *number* in $3 + n = 5$. This technique of selecting letters to suggest what they stand for has strong pedagogical appeal and is undoubtedly helpful to beginning students. However, Peter Rosnick (1981) conjectures that algebra students may be seduced into making certain kinds of errors by thinking of, say, a as representing apples, when, in fact, a properly represents the *number* of apples.

A third similarity between letters and words is that they can both mean different things in different contexts. That is, the meaning of the word *run* and the meaning of the letter x both depend on the contexts in which they are used. (The *meaning* of a literal symbol is derived from its *role* [as name, unknown, indeterminate, parameter, etc.], its *domain* of values, and its associated *truth set* [if used in an open sentence].) This similarity is often taken for granted and may seem significant only when contrasted with the differences.

Differences. Although letters and words are similar in that they may assume different meanings in different contexts, they differ with regard to their consistency of meaning throughout a single context. According to mathematical convention, the meaning and, in particular, the value ascribed to a literal symbol must be the same wherever that symbol appears in a given context. That is, in substituting values for x in the sentence $3(x + 2) + 5 = 17 - 2x$, the same value must be substituted wherever x occurs.

This is not true for verbal expressions. Identical words or phrases may refer to different things within a single sentence. Colloquial examples of this phenomenon are easy to generate: "His *leg* became numb during the last *leg* of the flight." But, even in mathematical contexts, the same verbal expression can refer to different things. Consider, for example, the proposition, "The sum of *an odd number* and an even number is always *an odd number*." Except for the instance in which zero is the even number, we all (somehow) know that the corresponding values of the first and second occurrences of the phrase *an odd number* are, in fact, different. It is only through a logically rigorous translation of this verbal proposition into mathematical language,

$$\forall x \forall y(x \text{ is odd } \land \ y \text{ is even} \rightarrow \exists z(z = x + y \land z \text{ is odd})),$$

that our tacit understanding becomes explicit and different literal symbols are used to represent each occurrence of the phrase *an odd number.*

The most significant differences between letters and words stem from letters not being associated with fixed sets of meanings the way words are. That is, we could never compile a dictionary of the meanings of literal symbols as they are used in mathematics. Instead, we are free to delimit the meaning and, in particular, the domain of most literal symbols in any way we wish—the *freedom of delimitation* property. Conversely, we are free to choose almost any literal symbol to represent a given referent—the *freedom of choice* property. (Mathematically, the *referent* for a literal symbol is its domain, or replacement set. Psychologically, the referent for a literal symbol in an open sentence tends to be its truth set.)

The *freedom of choice* property of literal symbols is what gives mathematical language so much flexibility.

The *freedom of delimitation* property of literal symbols is what gives mathematical language so much *generality.* Because literal symbols can be used to represent any set whatsoever, a single statement can, once and for all, define a concept or assert a relationship. Changes in the referent do not necessitate changes in the literal symbol used to represent the referent.

As for verbal expressions, a given word or phrase can automatically restrict the referent set and thereby limit the generality of a statement. For example, in the sentence, "*He* is a mathematics teacher," the referent set for the pronoun *he* is implicitly restricted to males. If the referent set is expanded to include females, the word *he* must be changed to an expression that can represent both males and females.

In contrast, the sentence, "x is a mathematics teacher," can refer not only to males and females, but even to computers. The literal symbol x imposes no inherent restriction on the generality of the statement and allows us the freedom to delimit the domain of discourse as we see fit.

Literal symbols grant the freedom of delimitation because letters, in and of themselves, are not saddled with the connotative baggage that verbal expressions convey. It is true, of course, that certain letters have acquired context-specific connotations through long traditions of usage. For example, in the absence of instructions to the contrary, we customarily consider x to be the independent variable and y to be the dependent variable in statements of functional relationships. We adhere to these conventions to save ourselves the trouble of delimiting the interpretation of each letter every time we use it. Nevertheless, we can, and often do, violate such traditions of usage by simply specifying the role and referent of a letter to suit our purposes.

The freedom of choice property of literal symbols is what gives mathematical language so much flexibility. That there are many freely interchangeable literal symbols to represent any given referent (infinitely many, if we count subscripted variables: a_1, a_2, a_3, …) means that changes in literal symbols do not necessarily imply changes in the referent. Thus we can substitute equivalent expressions, as in the composition of functions, to deduce logical relationships among variables without altering either the actual or implied meaning of the literal symbols. (If $y = 2x + 3$ and $z = 5y - 7$, then $z = 5(2x + 3) - 7$.)

With verbal expressions, the nature of the elements in a set definitely limits the words and phrases that can be used to refer to that set. Consequently, changing a verbal expression nearly always implies some change in referent. For instance, changing *he* or *she* in the sentence about mathematics teachers changes the referent set from males to females.

Of course, not all changes in verbal expressions are quite so dramatic as changing *he* to *she*, but even supposedly synonomous words and phrases imply subtle changes in referent. That is, *he* is not the same as *this man* who, in turn, is not the same as *that man*. Examples of words and phrases that mean *exactly* the same thing are rare indeed.

In contrast, A, or T, or almost any other letter (except I, perhaps) would do just as well as x to symbolize the elements in a set of mathematics teachers. Here again, we often follow certain conventions in choosing literal symbols in order to save time and to give students a certain sense of consistency. Nevertheless, we always have the flexibility of contravening traditions of usage whenever we find it useful to do so.

Teaching Suggestions

How can we, as teachers, help students develop a better understanding of literal symbols? First, we ourselves need to be aware of the myriad ways these symbols are used and recognize the particular characteristics they exhibit in various contexts. Next, we need to alert students to the properties of literal symbols and point out which characteristics may be similar to words or numerals and which characteristics are unique to the literal symbols themselves.

If we want students to gain a real appreciation for the power of literal symbols and yet not be overwhelmed by a lengthy, and probably meaningless, discourse on their various characteristics, then we need to introduce these ideas gradually as different uses of literal symbols appear in the curriculum. For example, in the early grades, when we use P to label a point or N to stand for a number, we can say that these letters are like abbreviations for words. Later, when we use arbitrary letters as labels, we can explain that these letters are like names for things. When we start using arbitrary letters in numerical contexts, we should mention that there is no connection between alphabetical order and numerical order.

As students move to the level of algebra, they should begin to realize that a letter behaves like a numeral in that it may represent a single number and may be subject to numerical operations and relations; they must also realize that a letter behaves very differently from a numeral with regard to the juxtaposition convention and the sign of the number. They should recognize that a letter is similar to a word in that it can mean different things in different contexts, but a letter is different from a word in that it must refer to the same thing throughout a single context.

Considering the placeholder analogy between words and literal symbols can help students appreciate that letters often represent many different numbers simultaneously. It is at this point that we can discuss the *freedom of delimitation* property of literal symbols, and we can contrast the generality of letters with the connotative richness of verbal expressions.

We should further point out to students that delimiting the domain of a literal symbol is not only a prerogative but also an obligation—an obligation that some textbook authors either overlook or fail to make explicit. We should assist students in determining the domains of literal symbols as they are used in textbook exercises and explain that an unspecified domain is generally assumed to be the real numbers.

The *freedom of choice* property of literal symbols seems to be an idea that students are quick to recite but slow to internalize. That is, many students who *say* that different letters can be used to represent the same thing often seem to *believe* that different letters must represent different things.

For example, some prealgebra and beginning algebra students seem to think that changing the literal unknown in an equation may also change the solution (Wagner 1981). We might check for this misconception by asking students to solve a given equation, then writing the same equation beside it with a different unknown and asking them for the solution to the second equation, without solving.

Later in the algebra course, when we give numerical illustrations of number properties, like the associative law or distributive law, we can use some examples in which different letters have the same value. Naturally, we want to emphasize the generality of number properties by picking arbitrary (and usually different) values for each letter in a statement. But, if we always avoid the special cases in which two or more of the letters have the same value, we may inadvertently reinforce the false notion that different letters represent different numbers.

In closing, I should mention that the advent of computers has presented students with a usage for literal symbols that often appears contradictory to the traditional mathematical usages—namely, the recursive use of letters as epitomized in the indexing statement $n = n + 1$. Research is currently under way to investigate the effect that computer programming experience may have on students' overall understanding of literal symbols.

Other current research is following the lead of Menger (1956), Collis (1975), and Küchemann (1978) in trying to catalog the various uses of literal symbols in mathematics. Obviously, it will be some time before we have a complete picture of what these things are that we call variables. Of course, as teachers, we cannot afford to wait for all the answers. We can, and should, start today to help students overcome the difficulties that have already been identified.

REFERENCES

Beberman, M. "An Emerging Program of Secondary School Mathematics." In *New Curricula*, edited by R. W. Heath. New York: Harper & Row, 1964.

Collis, K. F. *The Development of Formal Reasoning* (Report of Social Science Research Council Project HR2434/1). Newcastle, New South Wales, Australia: University of Newcastle, May 1975.

Herscovics, Nicolas, and Carolyn Kieran. "Constructing Meaning for the Concept of Equation." *Mathematics Teacher* 73 (November 1980):572–80.

Hockett, C. F. *Language, Mathematics, and Linguistics.* The Hague: Mouton, 1967.

Küchemann, D. "Children's Understanding of Numerical Variables." *Mathematics in School* 7 (1978):23–26.

Matz, M. *Towards a Process Model for High School Algebra Errors* (Working Paper 181). Unpublished manuscript, M.I.T. Artificial Intelligence Laboratory, 1979.

Menger, K. "What Are *x* and *y*?" *The Mathematical Gazette* 40 (1956):246–55.

Rosnick, Peter. "Some Misconceptions Concerning the Concept of Variable." *Mathematics Teacher* 74 (September 1981):418–20, 450.

Russell, B. *An Inquiry into Meaning and Truth.* London: Allen & Unwin, 1940.

Tarski, A. *Introduction to Logic.* New York: Oxford University Press, 1941.

Wagner, Sigrid. "Conservation of Equation and Function under Transformations of Variable." *Journal for Research in Mathematics Education* 12 (March 1981):107–18. ◆