

An investigation into the modelling of word problems leading to algebraic equations

Ruth Hubbard

Queensland university of Technology

Abstract: The difficulties experienced and the errors made by students at all levels, from elementary school to university, with turning word problems into mathematical statements have been investigated from many points of view. The main aim of the present study, using first year university students as subjects, was to discover whether there are any differences in the success rates of students between problems in which they just translate the word problem into an algebraic equation and problems in which they are asked to solve the equation as well. Although the success rates were not very different, a large proportion of the correct solutions to the equation were obtained without using an algebraic equation as part of the solution process. This suggests that it might be helpful to give students opportunities to model equations from word problems in which a solution is not possible because there is no numerical information provided, before they are asked to solve equations. A secondary aim of the study, motivated by the Student/Professor problem, was to investigate students' errors when the relationship between two variables was given as a quotient instead of a product.

Introduction

The errors students make when attempting to write simple linear algebraic equations have been studied for many years and by a large number of researchers, an early example being Paige and Simon (1966). Nevertheless some of the mechanisms involved in the translation process, which cause the errors, remain elusive. Of course the modelling process depends on an understanding of the meaning of algebraic variables and this has also been a focus of much research, for example Kuchemann (1978). More recently MacGregor and Stacey (1997) have shown that students' misunderstandings about the meaning of algebraic symbols, depends not only on cognitive development but on various environmental factors, one being the type instruction used to introduce algebraic symbols.

Two processes for translating word problems into algebraic equations, which lead to errors, syntactic translation and static comparison were identified by Clement (1982).

The process of syntactic translation, where students use the word order in the story to form the equation, frequently results in the relationship between the two variables being reversed, as in the now famous students and professors problem,

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." Write S for the number of students and P for the number of professors.

The answer, $6S = P$ is claimed to be an example of syntactic translation because the word "six" directly precedes the word "students". On the other hand, static comparison produces the result $6S = P$ because there are more students and the number six expresses this fact. However Macgregor and Stacey (1993) showed very convincingly, by using questions in which syntactic translation would have produced the correct equation, that many students still reversed the variables in their responses. They therefore concluded that the cognitive models that students use to write mathematical equations are not linear but very complex and that a great deal more research is required on how students use these models to construct mathematical equations and how they can be helped to develop the thought processes required to translate word problems into equations correctly.

The present study further investigates some aspects of this problem. It compares students' responses to two problems, which have similar syntax, it discusses the importance of the particular letters used in the statement of the problem and investigates the methods students use to solve an equation. Because most students complied with the request to explain their responses, a number of other issues emerged from an analysis of their responses, which were not a part of the original design. In studies where students just write an answer or select one from the answers

provided this additional information is not available. More detailed information then needs to be obtained from interviews.

Even though the research into the modelling and solution of simple algebraic equations has produced results that could well lead to improved teaching strategies this does not appear to have occurred. The problems identified twenty years ago are still there, both at school and undergraduate level. Some further suggestions for improving the teaching of elementary algebra will be made.

Research questions

1. Are students more successful at translating a statement given in words that leads to an expression or to an equation?
2. How do students respond to translating a statement involving a quotient rather than a product?
3. How are students' responses to modelling word problems affected when a numerical solution is asked for?

Method The students in this study were a class of first year engineering students at the Queensland University of Technology. There are two first year mathematics courses for engineering students. One class of 243 students, consisted of those who were considered to be well prepared for university mathematics and one class of 232 students consisted of those who were less well prepared but who had studied mathematics to the end of secondary school. The second group were chosen for this study. All these students would have been at least 17 years old, many considerably older.

The students were asked to answer the following three questions in their second week at university.

1. In a QUT classroom, there are 2 chairs beside each table. If there are n chairs, how many chairs and tables are there altogether?
2. In an Engineering Maths test, the number of students who pass, is 3 times the number of students who fail. If the number of students who pass is n , write an equation for t , the total number of students in the class.
3. In an Engineering tutorial, there are twice as many females as males. (A very unusual occurrence.) If there are 33 students in the class and x of them are females, how many males are there in this class?

Technically, all the letters required for making correct responses are provided, but students were asked to label any new letters they wished to introduce. As stated, the questions all involve quotients rather than the products used by most previous researchers. This was intended to make the questions a little more difficult because the sample consisted of university, rather than junior secondary school students. In several pilot studies, questions of varying difficulty were tried but it became evident that if the questions were more difficult than those above, a large proportion of students would simply make no attempt to respond. The questions were not part of any formal assessment so the students had to feel that they were able to make a response without being coerced into doing so. In all, 217 responses were obtained, which it is reasonable to assume, was the number of students who were present at that particular session.

In the first and second questions, the syntax is very similar but the first asks for an expression while the second asks for an equation connecting two variables. Neither of these questions provides enough information for obtaining a numerical solution. The third question provides more information and does ask for a solution. There was a large space below each question and students were asked to write any working, explanation and thinking steps.

Results

Of the 217 responses, 56 answered all three questions correctly and were excluded from most of the remainder of analysis. The questions were clearly too easy to challenge these students. These 56 students were only included in the part of the study which looked at students' use of their own

variables. The questions themselves contained all the variables **necessary** to write the response and these variables did not use the first letters of any of the words in the problem. **Table 1.** shows the proportions of students who introduced variables which were the first letters of the objects being counted in at least one question. If students introduced another variable such as y , which was not the first letter of a key word, this was ignored.

	All 3 correct (N =56)		Not all correct (N=161)		Total	
	No.	%	No	%	No	%
Initial letter as variable name	16	29	75	47	91	42

Table 1.

Sometimes the student did not introduce a new variable but felt the need to rename an existing variable, to reflect the object it was counting, for example in Q3. $x = f$, in other cases the new variables were introduced for the other objects named in the questions, tables, failing students and males. The table suggests that the weaker students, those that did not get all three questions correct, tended to need letters that represented “things” more than the better students. The use of letters that bore no visible relationship to the objects they were supposed to be counting, appeared to add another level of abstraction to the problem. Most students complied with the instruction to label any new variables they wished to use but statements like “ $t = \text{tables}$, $c = \text{chairs}$ ” were very common and in some cases that is exactly what the students meant because such statements were then followed by formulae such as $T = 2cn + 1tn$ or even $A = (2c + 1t)^n$. This type of error, where n is used as an adjective, qualifying the name of the object being counted, occurred 13 times. This phenomenon has been observed in other studies, for example Stacey and MacGregor (1997) observed that this phenomenon persisted when junior secondary students were tested in three years. These results show that for some students this misconception persists much longer.

Expressions versus equations

The only essential difference between questions 1. and 2., is that the first question asks the student to write an expression while the second asks for an equation connecting two variables. Mathematically an expression is a simpler object than an equation because no relationship is implied and an understanding of equality is not required. However students are clearly accustomed to being required to write equations so many simply inserted the letter t and an equal sign. Provided the expression was correct, this was regarded as a correct response. The success rates of students in these two questions are shown below.

Question 1 (Expression)	Question 2 (Equation)
Correct	Correct
44%	60%

Table 2.

This is the opposite of the expected result and my only explanation is that in local beginning algebra textbooks, students’ first encounter with algebraic symbols is with writing functional relationships. This is followed later with modelling simple equations, so they may have had no experience in writing expressions.

Inconsistency

The wording of the first two questions was intended to be very similar, with just the symbol for total absent in Q1. This enabled a comparison to be made between students who modelled both questions as quotients, both as products or who wrote one of each. Table 3. shows the frequencies of these pairs of responses.

	Both products	Both quotients	Mixed	Unclassified
Number	24	80	42	15
Percentage	15	50	26	9

Table 3.

A response was labelled a product if it had the form $t = n + 3n$ or $3n + f = t$ or if it was not quite complete but the product form was clear. A typical quotient response was

$$n + \frac{n}{2} \text{ or } f = \text{no. who failed, } t = 3f + f, \text{ where the quotient has been turned into a product but}$$

is a correct translation of the information. The unclassified responses were the result of both questions not being answered, of equations involving other operations such as differences or powers or illegible responses. Such a large number of inconsistent responses suggests that some students have conflicting processes, whether they be syntactic translation or static comparison or some other procedure, for dealing with this kind of problem and that the particular response that is made depends on which process gets the upper hand at any particular moment. This kind of vacillation was evident in the student interviews reported by Clement (1982) in connection with the students/professors problem. Students might write the correct equation, then decide it was not correct and replace it with an incorrect one or vice versa. This was also happening in some of my written responses where students wrote an equation, crossed it out and replaced it with a different one, usually quotient \rightarrow product or product \rightarrow quotient. In these cases the student's final uncrossed-out response was used but it was clearly not very reliable.

Solving the equation

When a problem requires a numerical solution, students focus almost exclusively on "getting the right answer". This is what has been emphasized in their mathematical upbringing from an early age and what is learned early in life is difficult to alter. As a result, when confronted by Q3, a large proportion of students abandoned algebraic symbols, which they had used in the first two questions, and concentrated on producing an answer.

	No attempt to write an equation	Wrote equation with an incorrect answer	Wrote equation and a correct answer	Unclassified
Number	59	23	57	22
Percentage	37	14	35	14

Table 4.

Most of the unclassified responses did not answer this question. Most of the students who wrote equations, introduced new variables, and at least initially, wrote an equation with two variables. The three equally popular variable pairs were x and m , x and y , or m and f . For some of these students the learned relationship between x and y took priority over the context related m and f . The reasons why some equations did not lead to a correct solution were that:

- once a second variable was introduced, the student could not eliminate it
- the equation contained a fraction, $x/2$ and the student could not solve such an equation
- the symbols representing the number of males and the number of females were used inconsistently.

Those students who did not attempt to write an equation, and showed their reasoning, used proportion, mostly successfully. Interestingly, when using proportion, the students did not reverse the variables, whereas those who attempted to write equations did so frequently, as in the first two questions.

Discussion

It is clear that before students can successfully translate word problems into expressions or equations they first need to have a thorough grasp of the meaning of algebraic symbols. The fact that 13 students demonstrated that for them, the letters represent things and that 91 out of the whole group of 217, or 42% needed to use the initial letters of words as an aid to forming equations, suggests that their understanding of algebraic symbols is not very thorough. This translation of symbols is frequently seen in reverse when students learn a mathematical procedure such as finding a derivative. If the function to be differentiated contains a variable other than x , then students whose understanding of the procedure is not very robust, change the variable to x first, and then apply the procedure.

The reason why students write an equation when the context only requires an expression, is probably because they have only encountered the translation into symbols in the context of equations. An examination of commonly used local school textbooks for the year in which algebra is introduced, shows that they do provide a few exercises in which students are asked to construct expressions from words or diagrams. However they then move on to looking at algebra in different ways, without making any attempt to integrate all these different approaches. The idea of translating words into formulas may never be revisited. Furthermore, the chapters on algebra are interspersed with chapters on geometry and various arithmetic techniques, so that if the sequence in the textbook is followed in instruction, it is difficult for students to make the necessary connections between concepts.

Although not originally research questions, the results showing the inconsistency of students' responses suggest that some of the hierarchies of understanding the meaning of variables and on reversal when writing equations, may not be as clear-cut as they first appeared. These students have had numerous mathematical experiences during their schooling and their problem appears to be that these experiences have not developed into a consistent cognitive structure. There is a conflict in students' minds between different interpretations and at any moment the answer to any single item is somewhat arbitrary.

Question 3. did not specifically ask students to use algebra to obtain a solution but it did include a variable and was preceded by two questions where algebraic symbols were necessary. So 49% followed these hints and used algebra but a surprising 37% chose to ignore the hints and go straight to an arithmetic procedure. A small number of students used arithmetic and then attempted to justify the solution algebraically. Again there is evidence of a conflict: do what the teacher wants and write an equation versus get the answer in the simplest possible way. Some students, who answered the first two questions correctly, wrote reversed equations in Q3. A possible explanation is that they confused the writing of the equation with the solution process. This was not clear and needs more careful investigation.

A personal perspective

I have come to this investigation from a background of preparing adults for university mathematics. My students range from those who are full of confidence about their mathematical ability but unknowingly bring with them a full complement of misconceptions, to those with serious maths anxiety who began to fail in mathematics in elementary school. My main aim is to draw together their varied experiences into one consistent cognitive structure into which their further studies at university can easily be incorporated. Anecdotal evidence from past students suggests that, to their great delight, that is exactly what happens.

We make the order of operations in arithmetic the basis of algebra, something against which many misconceptions can be tested. We begin modelling expressions from word problems in the first lesson, by taking first a story with numbers followed by a similar story with letters and building up the complexity of the expressions until they contain many different operations. We use counterexamples constantly to make all the misconceptions explicit. Some habits such as "guess and check" to solve equations, are difficult to dislodge because they have proved successful for many years. Because of this, I think "quick-fix" methods for solving all kinds of problems should be avoided, and more general, robust methods should be developed from the very beginning. With young students it has to be done more slowly, with constant reinforcement but I am sure it can be done.

References

- Clement, J.:1982, 'Algebra word problem solutions: Thought processes underlying a common misconception', *Journal for Research in Mathematics Education* **13**(1), 16-30.
- Kuchemann, D.:1978, 'Children's understanding of numerical variables', *Mathematics in Schools*, **7**(4), 23-27.
- MacGregor, M. and Stacey, K.:1993, 'Cognitive models underlying students' formulation of simple linear equations', *Journal for Research in Mathematics Education* **24**(3), 217-232.
- MacGregor, M. and Stacey, K.:1997, 'Students' understanding of Algebraic notation' *Educational Studies in Mathematics*, **33**, 1-19.
- Paige, J.M. and Simon, H.A.:1966, 'Cognitive processes in solving algebra word problems', in B. Kleinmuntz (rd.), *Problem Solving: Research, Method and Theory*, Wiley, New York, pp. 51-148.