Mollie MacGregor & Kaye Stacey

STUDENTS' UNDERSTANDING OF ALGEBRAIC NOTATION: 11-15

running head: Understanding algebraic notation

ABSTRACT

Research studies have found that the majority of students up to age 15 seem to be unable to interpret algebraic letters as generalised numbers or even as specific unknowns. Instead, they ignore the letters, replace them with numerical values, or regard them as shorthand names. The principal explanation given in the literature has been a general link to levels of cognitive development. In this paper we trace the origins of specific errors to four additional factors that may or may not be associated with cognitive levels. We present evidence that some of these errors reflect the use of intuitive assumptions and pragmatic reasoning, analogies with familiar symbol systems, interference from new learning in mathematics, and the effects of poorly-designed and misleading teaching materials.

STUDENTS' UNDERSTANDING OF ALGEBRAIC NOTATION: 11-15

running head: Understanding algebraic notation

The *Concepts in Secondary Mathematics and Science* [CSMS] research project (Hart, 1981) provided evidence linking students' levels understanding of algebraic letters to Piagetian stages of cognitive development and to IQ scores. It was concluded that most of the 13-15-year-olds tested were unable to cope with items that required interpreting letters as generalised numbers or even as specific unknowns. In the many years since the CSMS project, it has been widely accepted that cognitive level is a sufficient explanation for the way in which algebraic notation is interpreted. If cognitive level is viewed as a barrier to the construction of certain concepts, it explains why students cannot do certain algebraic tasks. However it does not explain why they interpret the notation in particular ways and why they make certain errors. In this paper, we explain some common misinterpretations by considering factors more accessible to diagnosis than cognitive level. We present evidence that difficulties in learning to use algebraic notation have several origins, including:

• intuitive assumptions and sensible, pragmatic reasoning about an unfamiliar notation system;

• analogies with symbol systems used in everyday life, in other parts of mathematics or in other school subjects;

• interference from new learning in mathematics;

• poorly-designed and misleading teaching materials.

The research reported in this paper is part of a larger project investigating the cognitive and linguistic demands of learning algebra in secondary school. In this project, data were obtained from pencil-and-paper tests given to a large representative sample of approximately 2000 students in years 7-10 (ages 11-15) in 24 Australian secondary schools. Some schools used the same test across two, three or four year levels, thus providing us with comparative data for these year levels. Other schools tested the same cohort of pupils on two or three occasions, thus providing us with longitudinal data on the progress of individual students. At one school we interviewed 14 students while they worked on selected items, audio-taping the discussions for later analysis.

The schools involved were not randomly selected. They took part in the tests because their teachers were keen to participate in the project and find out about the effectiveness of their teaching. However, because of the sample size, the number of schools, and the range of school types (State, Catholic and private, in working-class and middle-class suburbs), there can be little doubt that those findings which are common to all schools in the sample apply to the general population of students. Results which are not uniform across schools point to the influence of factors specific to particular schools.

BACKGROUND

Students' Interpretations of Algebraic Letters

In Australian schools, students begin algebra in Year 7 or Year 8 when they are 11-13 years old. In the first year, students are taught to use letters to stand for unknown or generalised numbers, frequently in the context of writing formulas for patterns. They are given opportunity to learn how to write simple expressions and equations containing letters, numerals, operation signs and brackets. Difficulties in learning these fundamental aspects of algebraic notation are well documented (Assessment of Performance Unit, 1985; Booth, 1984; Cambridge Institute of Education, 1985; Herscovics, 1989; Küchemann, 1981; Robitaille &

Garden, 1989).

Küchemann (1981) classified students' interpretations of algebraic letters into two major divisions:

1. Student ignores the letter, gives it an arbitrary value, or uses it as the name of an object.
2. Letter is used as specific unknown number or as generalised number.

Each of these divisions is further divided into two categories to account for the cognitive demands of item complexity, thus giving four levels. Küchemann suggested that these four levels correspond to the Piagetian stages of *below late concrete*, *late concrete*, *early formal* and *late formal*. Longitudinal testing in the CSMS project showed that students with higher scores on a non-verbal IQ test tended to demonstrate higher cognitive levels and made faster progress through levels than students with lower IQ scores. Nevertheless the fact that, in algebra, a few students with below-average IQ scores reached the third or fourth levels by age 15 (see Hart, 1981, p. 185, Fig. 12.4) suggests that other factors need to be taken into account when explaining students' growth of understanding of algebra. In the present paper we indicate what these other factors might be.

Approaches to Beginning Algebra

The pattern approach to algebra, currently presented in curriculum advice to schools (Stacey & MacGregor, in press), depends on students' ability to grasp the concepts of generalised number and unclosed expression. In this approach, students' first use of algebraic notation is for expressing relationships in patterns and sequences, where there are two variables related by a rule (e.g., $y = 2x + 1$). The fact that this approach is widely recommended by mathematics educators as a gateway to algebra stands counter to the argument that the concepts involved and the notation for expressing them are beyond the capacity of ordinary students. However capacities should not be considered outside a context for their expression. The success of students in computer environments (Cohors-Fresenborg, 1993; Sutherland, 1991; Tall & Thomas, 1991) suggests that at least some of the difficulties and errors in traditional algebra learning are caused by the nature of students' learning experiences and do not reflect their cognitive capacities.

RESEARCH QUESTIONS

Küchemann (1981) concluded that the majority of 13-15-year-olds are unable to cope with algebraic letters as unknowns or generalised numbers. If this is the case, then current approaches to algebra as a language for expressing relationships between two variables, whether via computer or with pencil and paper, are not appropriate. In the current climate of new approaches to algebra instruction, it is important to reassess students' capabilities. In particular, we wanted to find out whether most beginners interpret letters as standing for numerical values or names of objects (Küchemann's lowest level) and to explore the roots of these interpretations.

In the testing programme, we assessed students' capabilities in several areas of algebra, including the recognition of operations and structures, the understanding of simple functions, and the ability to construct and solve equations. In this paper we report results for only a few items in the tests - those concerned with the interpretation of algebraic letters and the writing of a simple unclosed expression. We discuss the following questions:

1. How do students who have not learned any algebra interpret

letters and try to write expressions? Do they come to algebra with preconceptions about the use of letters?

2. How do students' interpretations of letters and simple algebraic expressions change over three years of school algebra learning?

3. What are the roots of specific errors and misunderstandings?

TESTING AND RESULTS

We discuss the testing and results in three sections. First we look at the way in which algebraic letters were interpreted by 11-12 year olds who had not been taught any algebra. We then comment on the progress made by these children in an eight-week algebra unit that formed part of their normal Year 7 curriculum. Next we report the results of tests used for several hundreds of students in Years 7 to 10 in 22 schools, with particular focus on their interpretations of algebraic letters and their ability to accept an unclosed answer. Interviews with individual students at another school provided us with insights into the reasons for certain errors in the main sample. Finally we trace the progress of 156 individual students in three schools who were tested three times: twice in one year and once the following year.

Year-7 Students' Progress in 8 Weeks

Pre- and post-test items

Items containing algebraic letters were included in a pre-test for two mixed-ability classes (n = 42) of Year 7 students (age approx. 11-12 years) who had not been taught any algebra at school. The same two classes were tested again eight weeks later after they had completed an introductory algebra unit. Two test items, shown in Figure 1, were used to assess students' ability to use an algebraic letter to represent an unknown quantity. Superficial differences were made to both items for the post-test, so that students would not be able to use remembered answers.

Pre-test	Post-test
1. David is 10 cm taller than Con. Con is *h* cm tall. What can you write for David's height?	1. Con is 8 cm taller than Kim. Kim is *y* cm tall. What can you write for Con's height? ....
2. Sue weighs 1 kg less than Chris. Chris weighs *y* kg. What can you write for Sue's weight?	2. Sam is *x* cm shorter than Eva. Eva is 95 cm tall. What can your write for Sam's height? ....

Fig. 1. Pre- and post-test versions of Items 1 and 2

Expected responses

We expected that in the pre-test most if not all of the students in our sample would not attempt the items containing algebraic letters because they had not been taught any algebra. If answers were written, we expected them to be at Küchemann's lowest level (i.e., letter ignored, given a numerical value, or used as a label for an object). A possible low-level interpretation of *h* in Item 1 (pre-test version, referred to as DAVID in this paper) is that it stands for the word "height". If "height" is denoted by *h*, then "David" should be denoted by D. We expected therefore that students might write D*h* to mean "David's height". In Item 2 (SUE), the letter *y* is clearly not an abbreviation for any word in the problem, and the only possible low-level response would be to ignore it or assign it a numerical value. We expected that many students would give no answer or a numerical answer.

Results

In the pre-test, two-thirds of the students did not write any answers, but the responses of the other 14 are useful indicators of students' intuitive interpretations of what algebraic letters might mean. Table I shows the responses to Item 1 and the likely explanations for them.

Table I

Responses to item DAVID from 14 students

Frequency	Response	Assumed reasoning
1	10+*h*[correct]	Add 10 to number or quantity denoted by *h.*
1	*h*10	Add 10 onto *h.*
1	U*h*	Abbreviated words "Unknown height".
2	18	*h* is the 8th letter of alphabet, therefore 10 more is the 18th.
2	110	Think of a reasonable height for Con, add 10
2	*t*, *g*	Choose another letter or adjacent letter for David's height.
5	10,20,"half"	No comprehension of the question; use of the given value 10 and operations "double" or "half".

We see in Table I nine sensible answers to what must have seemed a strange question. Two students used *h* to represent a quantity to which 10 cm could be added, and one of them wrote the correct expression 10 + *h*. One student used letters as abbreviated words. Two students associated the letter *h* with its position in the alphabet, as they often have to do in puzzles and translation into codes. Two students thought that they should assume a value for Con's height, since it was not given. Two students reasoned that if Con's height could be represented by a letter, then so could David's height. The remaining five students, not knowing what to do, had tried to write something related to the numbers in the question, ignoring the letter. Responses to Item 2 by the same students showed that they were almost all reasoning consistently. Table II shows the responses of these 14 individuals to the two items, and explanations for their responses to Item 2 (SUE).

Table II

Responses to items DAVID and SUE from 14 students

Frequency	DAVID	SUE (assumed reasoning)
1	10+*h*	*y*-1 (Subtract 1 from number or quantity denoted by *y*)
1	*h*10	*x* (Although 10 can be "joined" *h*, as 10*h*, 1 cannot be "removed" from *y*. To denote 1 less than *y*, write *x*)
1	U*h*	U*w* (to mean "Unknown weight")
2	18	24 (*y* is the 25th letter, therefore 1 less is 24)
2	110	[no response]
2	*t*, *g*	*o*, *x* (Choose another letter or adjacent letter)
4	10, 20	1 (No comprehension of the question; use of the given value 10 and operation "double")
1	"half"	[no response]

Interpretations of algebraic letters by these Year 7 students can be divided into six categories, listed below. The first three in the list correspond to Küchemann's lowest level of interpretation. The last in the list corresponds to Kuchemann's next level of interpretation (i.e., specific unknown or generalised number).

- *letter ignored*
- *numerical value*
- *abbreviated word*
- *alphabetical value*
- use *different letter* for each unknown
- *unknown quantity*

As described below, not all these categories were evident in the post-test.

The post-test was given eight weeks later, after the class had been taught their first algebra unit lasting about twenty lessons. In this test, given to 38 of the original sample, most students responded and many were correct (see Table III). Ten students wrote the terms 8*y*, *y*8, 8-*y* or *y*-8 for Con's height (see Fig. 1), in which their errors are due to conjoining terms for addition or writing the wrong operation. Combining these responses with the 14 correct ones, we see that 24 students (i.e., 63% of the sample) had understood the significance of algebraic letters well enough to write an expression containing the symbols *y* and 8. No student used a letter as an abbreviated word or person's name or assigned an alphabetical value, and only three students chose numerical values. Three chose other letters (*x* or N), these being letters they had seen used for algebra in their lessons. There were still a few students who had apparently not understood the question and wrote "8" (the numeral given in the item) or did not respond. Similar results were obtained for Item 2, with 16 correct responses and seven students using letters as unknowns but making syntax or operation errors (e.g., x - 95, -x95, or x + 95).

The performance of this class was impressive, and indicates the effectiveness of the teaching program that had been used. The majority appear to have learned that letters can stand for unknown numbers or quantities, and many were able to write an unclosed expression. We do not know, however, to what depth they understood the significance of letters in algebra. As we will see in the next section, older and more experienced students often misinterpreted algebraic letters in various ways and were not all able to provide unclosed answers.

Table III

Percentage of responses from Year-7 students in each of six categories before and *after instruction*

Category	Percentage of sample[a]	
	Pre-test (*n*=42)	Post-test (*n*=38)
unknown quantity [correct]	2%	37%
unknown quantity [incorrect]	2%	26%
abbreviated word	2%	0
alphabetical value	5%	0
numerical value	5%	8%
use of different letter	5%	8%
letter ignored	12%	10%
no response	67%	11%

[a]In the pretest, 14 of 42 students responded to both items. In the post-test, 34 of 38 students responded to both items.

Use of Algebra Notation at Four Year Levels

The items DAVID and SUE that had been used for the Year 7 sample were included in a test used by 22 schools for Years 7 to 10. All students in this sample had been taught some algebra. We assumed that students at all levels would have seen or written expressions like TWO OPERATIONS (referred to below as TWO OPS) in their early algebra experiences with number patterns, function machines, and "guess my number" rules, as well as in purely symbolic translation exercises. The older students in the sample would have seen algebraic letters used in a geometrical context, as in DISTANCE. The results for the four items are shown in Table IV.

3. What is the distance around these shapes?

(i) .... (ii) ....

4. *n* stands for an unknown number.
Write the following in mathematical symbols:
"Add 5 to *n*, then multiply by 3" ....

Fig. 2. Items 3 and 4, referred to as DISTANCE and TWO OPS in this paper

The Year 10 students were more successful than the Year 7 students, but there was not the great improvement that we had expected. As Table IV shows, only 73% of the Year-10 students were correct for the easiest item. Despite gains from one year to the next, the success rate for Item 4, the hardest item, did not reach 50%. As we explain below, errors due to several causes arise as students move from Year 7 to Year 10. In the following sections we describe the main errors and misunderstandings for each item and we discuss their causes.

Table IV

Percentages of students in Years 7-10 correct for Items 1-4 (*N* = 1463)

Item [answer]	Yr 7 (n=307)	Yr 8 (n=511)	Yr 9 (n=338)	Yr 10 (n=307)
1. DAVID [h+10]	39%	52%	63%	73%
2. SUE [y-1]	36%	46%	60%	64%
3(i) DISTANCE [3x]	42%	44%	65%	61%
(ii) DISTANCE [2x+18]	27%	35%	55%	53%
4. TWO OPS [3(n+5)]	14%	17%	25%	47%

Errors in items 1 and 2 at four year levels

Interpretation of letters. Success rates on DAVID and SUE at Year 7 were approximately the same as the rates for the previous Year 7 sample on their post-test, supporting our previous finding that about one-third of beginners are capable of learning to use a letter to represent an unknown quantity and to write an unclosed expression. Approximately half the Year-8 students and two-thirds of older students were correct for these very simple items. The incorrect responses reveal that letters were used with several different meanings. The intuitive interpretations made by Year 7 beginners were seen at all year levels. They were: *abbreviated words* ; *alphabetical value* ; *arbitrary numerical value; letter ignored* . New misinterpretations were:

- *label* associated with an object or quantity (e.g., C to mean "Con's height" and D to mean "David's height" in C+10 = D).
- *letter equals 1* unless otherwise specified (e.g., 10+h = 11)
- letter has *a general referent* that includes various specifics (*h*

means "height", so it means both "David's height" and "Con's height" in the statement h = h+10)

Numerical responses. Younger students often ignored the algebraic letter and chose a number for Con's height in Item 1. This response had also been found in the small sample of students who had not been taught any algebra, as discussed above. For Item 3 many Year-7 and Year-8 students measured the lengths marked "x cm" with their rulers. They probably thought this was what the teacher wanted. Our data indicate that some numerical responses from older students were not due to measuring, *letter ignored* or *arbitrary numeral*, or to the *letter* = *1* belief (explained in a later section of this paper), but were in fact the results of attempts to use algebraic manipulation to "solve an equation" without understanding. An example is the following, written by a student in Year 10 and producing the answer 5 for DAVID:

$$10 + h\frac{10 + h}{h} \infty \frac{h}{h}$$

$$= \frac{2h}{2} + \frac{10}{2}$$

$$h = 5$$

This student has written the correct expression 10 + h but has then tried to use routine manipulation techniques. It is likely that recent learning of procedures for simplifying algebraic fractions or solving equations has caused the retrieval of schemas related to that learning. It is possible that, in research studies reported in the literature, responses of this type as well as responses in the category *letter* = *1* have been misclassified as *letter ignored* or *arbitrary numeral*.

Conjoining. Studies reported in the literature found that conjoining for addition (e.g. 10*h* to represent 10 plus *h*) is a common error (see, for example, APU, 1985; Booth, 1984). In the present study, the incidence of conjoining for addition was far lower than expected at all levels, being well below 10% after Year 7. Moreover, since about half the students who wrote 10*h* for "10 plus *h*" wrote 1*y* for "1 less than *y*", what looks superficially like the conjoining error was likely to be merely a thoughtless combination of the numbers and letters presented in the question without concern for representing the operations linking them. A few other "conjoiners" appeared to believe that if a coefficient is on the left of the letter it indicates subtraction and if it is on the right it indicates addition. They wrote *h*10 to mean "add 10 to h" and 1*y* to mean "take 1 from y". We assume that this notion comes from their experience with adding and subtracting along the number line or from their knowledge of Roman numerals, or that it is a reasonable invented notation (referred to later in this paper).

<u>Errors in Item 3 at four year levels</u>

<u>Exponential notation.</u> The misuse of exponential notation (x^3 instead of $3x$) in this item and in other items in the test increased steadily over the four year levels, from 5% at Year 7 to 18% at Year 10. Our evidence, based on written responses as well as interview data, suggests that one major cause is lack of clear concepts for repeated addition, multiplication and repeated multiplication. (See Stacey & MacGregor, 1994, for evidence of students' confusion about these concepts and their notation.) Teachers may wrongly assume that students have a firm understanding of these concepts when exponential notation is taught and used.

<u>Interference from inappropriate schemas.</u> There appear to be many factors contributing to the difficulty of Item 3 which are not

specifically associated with conventions of notation or the meanings of letters. Older students had more opportunities for making mistakes than younger ones because of interference from new schemas only partly learned or because of their expectations of being able to use more advanced knowledge. When students were interviewed on this task, they made comments such as "Do I have to figure out the numbers?", "That's the hypotenuse", or "If it's 8 across that way, then rule off the line and cut straight up". Comments such as these indicate that they were searching for remembered schemas or learned procedures from geometry, or did not recognise that the task simply required a sum of measures. It was interesting to see that for part (ii) several Year 10 students wrote $x^2 + 5^2 + 8$. Their uncertainty about how to write "twice x" or "x plus x" may have contaminated their knowledge of how to write "twice five". Alternatively, they may have tried to recall knowledge about right-angled triangles and Pythagoras's theorem, as indicated by their questions and comments quoted above.

Interference from learned misinterpretation. Another obstacle to success in Item 3 is the belief that any letter stands for 1, evident in at least three of the schools in our sample. The origins of this belief became clear in the interviews. A Year-10 student said, "x is just like 1, like having one number". Another said "By itself it is 1, the x". For DISTANCE (ii) a student worked out the answer as 20, and explained that 8 plus two 5's is 18, then "1 more for each x makes 20". Explanations such as these enabled us to understand the reasons for many numerical answers to items in the written test. Answers that we had first classified as arbitrary numerical value (e.g., David's height = 11) or inaccurate measurement (e.g., 20 cm for DISTANCE (ii)), could in many cases be attributed to the *letter equals 1* belief. One likely cause of this belief is a misunderstanding of what teachers mean when they say "x without a coefficient means $1x$". The student gets a vague message that the letter x by itself is something to do with 1. Another cause of misunderstanding is the fact that the power of x is 1 if no index is written (i.e., $x = x^1$). Like numerical responses caused by misguided attempts to simplify or solve (see previous section), numerical responses caused by the *letter equals 1* belief may have been misclassified in previous studies as *letter ignored* or *arbitrary numeral*.

<u>Errors in Item 4 at four year levels</u>

In Item 4, which required students to coordinate two operations, most students at all levels were not correct (see Table 3). The low success rate is in accordance with the findings of Ursini (1990) for a similar item used for students in Mexican schools. In Table V the category "Omit brackets" includes all responses that would have been correct if brackets had been inserted, for example $n + 5 \infty 3$. We had expected that omission of brackets would account for most errors, but this was not the case, as Table V shows. The large percentage of unclassified responses (various combinations of the symbols 3, n and 5) and omissions points to students' lack of experience in using algebraic notation to express information.

We had expected conjoining to be a common error. It would cause $5n$ to be written for $5 + n$, and finally $15n$ when the $5n$ was multiplied by 3. However no more than 14% of students at any level wrote a conjoined answer. This figure is in accord with the relatively low incidence of conjoining in other items. A few students clearly needed an extra symbol for the result of the first operation (i.e., adding 5 to n) and they used another variable name (e.g., y) or left a space for the "answer". There were instances of the "unknown letter equals 1" belief, which gives the answer 18 when the student adds 1 to 5 and then multiplies by 3. Some students tried to write an equation. We have no explanations for students' reasoning behind many other forms of responses and the reluctance of so many to write anything at all.

Table 5

Results for Item 4 (N = 1806)

Year	n	Correct	Omit brackets	Use of exponent	15n	5n∞3	Other	Omit
7	368	14%	14%	2%	8%	6%	20%	36%
8	594	19%	22%	1%	6%	2%	26%	24%
9	386	26%	21%	2%	10%	4%	22%	15%
10	458	47%	11%	1%	8%	4%	9%	20%

Trends in Three Schools over 13 Months

Teachers at three schools tested their students three times to assess progress and locate persistent errors. At two of these schools students were tested twice in Year 8 and once the following year, Year 9. At the other school, students were tested twice in Year 9 and once the following year, Year 10. The items DAVID, SUE and TWO OPS were included in each of the tests, with superficial changes between tests (e.g., as shown in Figure 1). Table VI shows results for the 156 students who did all three tests.

Table VI

Percentages correct in three school groups each tested three times (n = 156)

	School A, Yr.8-9 (*n* = 70)			School B, Yr.8-9 (*n* = 60)			School C, Yr.9-10 (*n* = 26)		
Item	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
DAVID	70	86	96	70	90	90	50	69	65
SUE	57	90	93	75	95	90	38	65	46
TWO OPS	20	41	74	35	35	55	31	81	62

No teaching strategies or learning materials were suggested to the teachers at the three schools, except in the case of School C for TWO OPS as discussed below. Trends in success rates on the other table entries are therefore due to the normal teaching that took place. As Table VI shows, students at Schools A and B made good progress in coping with algebraic letters and unclosed expressions (DAVID and SUE). On the other hand, at School C improvement on these two items was small; the cause of this poor performance is discussed below. Success on TWO OPS improved at all schools, with a dramatic rise in the performance at School C after the teaching intervention.

<u>Responses to DAVID and SUE at School C</u>

Misinterpretation of algebraic letters was a persistent difficulty at this school over all year levels. At Year 9, several students wrote R for DAVID (i.e., ten letters after H) and X for SUE (i.e., one letter before Y), indicating their belief that the value of a letter is determined by its position in the alphabet. This error was easily corrected by the teachers after it was brought to their notice. Other misinterpretations, in particular the belief that "Con's height" should be represented by C or C*h* but not by *h*, were more resilient. Discussion with the teachers revealed that teaching materials that had been used in Years 8 and 9 for these students explicitly present letters as abbreviated words (e.g., *c* could stand for "cat", so 5*c* could mean "five cats"). In contrast, teaching materials used at Schools A and B present letters as standing for unknown numbers. In the data from these two schools, there were only two instances of letters used as abbreviated words in the first test and none thereafter. This observation suggests that the "letter as abbreviated word" is not an inevitable naive misinterpretation that must be corrected but is caused by certain teaching materials or teachers' explanations. It seems probable that widespread and persistent misinterpretation by the School C students can be attributed to the misguided teaching approach that had been used.

<u>Responses to TWO OPERATIONS at School C</u>

At School C, between the first and second tests teachers used a lesson designed to address difficulties in coordinating two operations as required in the item TWO OPS. This lesson used examples of English text as well as mathematics to make students aware of the potential for ambiguity in certain expressions and the support of context in English that is not present in mathematics. For example, in the phrase "French language and literature" the word "French" modifies both "language" and "literature". However in the phrase "French fries and coke" the word "French" is not a modifier of the word "coke". In the phrase "Twice five plus three" it is not clear whether "twice" modifies 5 only, or both 5 and 3. Students were given practice at generating expressions of this type, inserting brackets to resolve ambiguity, and evaluating them.

The teachers' efforts to teach the use of brackets for grouping and distributivity were clearly very effective. As shown in Table VI, the success rate rose from 31% correct in the first test to 81% correct in the second test. In the third test, given the following year, 62% were correct. In this third test, several students omitted brackets from otherwise correct expressions, suggesting that their knowledge of the purpose of brackets had not been used and was consequently forgotten. There were however no other types of error, in contrast with the great variety of errors seen in the first test. We conclude that the lesson on brackets and ambiguity had been effective for the majority of students. We are not sure why it also seems to have been effective in eliminating the other types of error. The reason may be that students had learned to focus more clearly on what an algebraic expression means and to see how a slight change in notation affects this meaning.

SUMMARY

Our findings show that important factors causing misunderstanding and errors in using algebraic notation from Year 7 onwards are:

- assumptions and guesses based on intuition or sensible reasoning and on analogies with other symbol systems;
- interference from new learning;
- particular learning materials used to introduce algebra.

We now summarise these findings and draw implications for teaching.

<u>Beginners' Assumptions about the Meanings of Letters</u>

Our results indicate that students' first attempts to interpret and use algebraic letters are based on sensible and pragmatic reasoning. We have shown that a Year 7 class who had learned no algebra guessed that letters stood for unknown quantities, abbreviated words or specific numbers. These specific numbers were either the alphabetical "value" of the letter or an arbitrary but reasonable value of the quantity described in the problem.

Certain initial misinterpretations of letters appear to be intuitive guesses, reflecting historical mathematical sign systems. For example, the interpretation of *a* as 1, *b* as 2, etc., has its parallel in the early Greek numeration system in which each number was denoted by a letter. Another example is the use of conjoining and ordering (e.g., 1y to mean "one less than y") which has its parallel in the Roman notation IV for "one less than five", VI for "one more than five", and so on. We suggest that in beginners these errors are not indicators of low levels of cognitive development; they are thoughtful attempts to make sense of a new notation or are caused by transfer of meanings from other contexts. In more experienced students, however, they indicate failure to learn, a condition which has many causes.

The success of the Year-7 students in our sample indicates that many of them were capable of learning to use letters as generalised numbers or unknown quantities and to accept "unclosed" expressions. They had moved beyond the stage of ignoring letters, assigning numerical values or using them as abbreviated names for objects. However at all year-levels there were some students who seemed to be unable to deal with the precise distinctions between symbols and their referents that is necessary for algebra. Many years ago, Paige and Simon (1966) described this as perhaps one of the most difficult aspects of using algebra to solve problems. Paige and Simon give an example of a problem about the value of a collection of dimes, for which students used x to stand for anything associated with dimes, e.g., " a dime", "the dimes", "the number of dimes" or "the value of the dimes". Similarly, in our items, some students in all age groups had difficulty distinguishing the name of an object (e.g., the person Con) from the name of an attribute (e.g., Con's height) and from a quantity or measure (*h* units). Wrongly interpreting the letter as the name of an object (e.g., interpreting *r* to mean "red pencils", so 6*r* means "six red pencils") is a well-known and serious obstacle to writing expressions and equations in certain contexts (see, for example, Clement, 1982; Kaput, 1987; for a deeper analysis of this difficulty and some new explanations see Lopez-Real, 1995; MacGregor & Stacey, 1993; Thomas, 1994). Moreover, students notice that concepts in applied mathematics are usually denoted by the initial letters of their names (A for area, *m* for mass, *t* for time, etc.). It is likely that this use of letters reinforces the belief that letters in mathematical expressions and formulas stand for words or objects rather than for numbers.

Effect of Teaching Materials on Meanings for Letters

Certain misconceptions about the meaning of algebraic letters were more prevalent in some schools than in others, and were traced to teaching materials that had been used. These misconceptions were the "alphabetical " interpretation ($a = 1$, $b = 2$, etc) and the use of letters as abbreviated words and labels (e.g., D*h* to mean "David's height").

The alphabetical interpretation was found to be common in one school at all year levels. Discussion with teachers at this school revealed that students had used many puzzles and games requiring this interpretation. Furthermore, students continually see letters used this way in textbook exercises labelled 1(a), 1(b), 1(c), etc., reinforcing their belief in a fixed value and order for each letter.

The persistent use of letters as abbreviated words and labels was traced to the use of textbooks that explicitly state, in the first algebra unit, that letters can be used as abbreviated names. There can be little doubt that the persistent misinterpretation of letters as abbreviated words, evident in data from schools using those books, was partly or wholly due to this misguided initial presentation.

Interference from New Learning

Our data show that after Year 7 a variety of misuse of algebra symbols arises as a result of interference by new learning. Exponential notation was used by many students to represent multiplication, and this incorrect use increased steadily over the four year levels. In a geometrical context, half-remembered area formulas and new learning about Pythagoras's theorem combined with uncertainty about the use of exponential notation to cause many errors.

We suggest that in a typical curriculum students do not get enough experience at using mathematical notation. In the schools we worked with, students learn algebra in one or two short modules per year. These modules are usually not connected with other work and have no useful purpose from the students' point of view. When algebraic concepts and methods are not used in other parts of the mathematics curriculum, students forget them and forget the notation for expressing them. They may remember certain surface features and spatial displays without the associated meaning and context. Consequently when new concepts and notation are introduced, students are unable to link these with, or differentiate them from, what they have previously been taught.

CONCLUSIONS

We have shown that students beginning algebra base their initial interpretations of letters and algebraic expressions on intuition and guessing, on analogies with other symbol systems they know, or on a false foundation created by misleading teaching materials. They are often unaware of the general consistency of mathematical notation and the power that this provides. Their misinterpretations lead to difficulties in making sense of algebra learning and may persist for several years if not recognised and corrected. The relative success achieved by some classes in the schools we tested and the poor performance by others suggest that factors such as different approaches to beginning algebra, teaching materials, teaching styles or the learning environment have a powerful effect. We have identified particular approaches and teaching materials that lead to misunderstanding and failure to learn. We have also shown that the majority of a group of 11-year-old students given appropriate instruction was capable of learning to write simple expressions in which algebraic letters stand for unknown quantities. On the other hand, there were students in other classes and schools whose initial experiences cause them to develop misinterpretations that in some cases persisted over four years.

In this paper we have identified several factors which we believe need to be considered in a comprehensive explanation of the particular ways in which students interpret algebraic letters. Attributing their interpretations to cognitive level alone does account for failure to write correct answers but does not explain the wide variety of misinterpretations that we observed in our data.

One of the general conclusions of the CSMS project was that any demand for abstraction in mathematics is beyond the capability of a large proportion of the secondary school population. It was suggested that, since the majority of students cannot cope with tasks where letters have to be interpreted as numbers, algebra teaching should be based initially at the level of concrete operations despite the fact that "the use of letters as objects totally conflicts with the eventual aim of using letters to represent numbers of objects" (Küchemann, 1981, p. 119). We suggest that this is a short-sighted approach, leading to "correct answers" to simple problems for the wrong reasons and reinforcing students' intuitive but unhelpful beliefs. As we have shown, misleading teaching materials in early lessons, intended to make initial algebra learning easy, can seriously disadvantage students.

We have shown that some common misunderstandings are the results of particular teaching approaches, and can be avoided; others have been developed by the students themselves, and when discovered can be overcome. Teachers need to be aware of the beliefs about the meanings of letters and mathematical notation that students bring with them to algebra learning, and take account of these beliefs in their teaching. They have a responsibility to ensure that students' first experiences of using letters for algebra lay the foundation for a coherent structure of algebraic knowledge.

REFERENCES

Assessment of Performance Unit [APU]: 1985, *A review of monitoring in mathematics 1978 to 1982*, HMSO, London.

Booth, L.: 1984, *Algebra: children's strategies and errors. A report of the Strategies and Errors in Secondary Mathematics Project*, NFER-Nelson, Windsor.

Cambridge Institute of Education: 1985, *New perspectives on the mathematics curriculum. An independent appraisal of the outcomes of the APU mathematics testing 1978-82*, HMSO, London.

Clement, J.: 1982, 'Algebra word problem solutions: thought processes underlying a common misconception', *Journal for Research in Mathematics Education*, **13**(1), 16-30.

Cohors-Fresenborg, E.: 1993, 'Integrating algorithmic and axiomatic ways of thinking in mathematics lessons in secondary schools', in *Proceedings of South East Asia Conference on Mathematics Education (SEACME-6) and the Seventh*

National Conference on Mathematics, Kampus Sukolilo, Surabaya, pp. 74-81.

Hart, K.: 1981, *Children's understanding of mathematics: 11-16*, Murray, London.

Herscovics, N.: 1989, 'Cognitive obstacles encountered in the learning of algebra', in S. Wagner and C. Kieran (eds.), *Research issues in the learning and teaching of algebra*, NCTM, Reston, pp. 60-86.

Kaput, J.: 1987, 'Towards a theory of symbol use in mathematics', in C. Janvier (ed.), *Problems of representation in the teaching and learning of mathematics*, Erlbaum, Hillsdale, pp. 159-195.

Küchemann, D.: 1981, 'Algebra', in K. Hart (ed.), *Children's understanding of mathematics: 11-16*, Murray, London, pp. 102-119.

Lopez-Real, F.: 1995, 'How important is the reversal error in algebra?', in B. Atweh and S. Flavel (eds.), *Proceedings of the 18th Annual Conference of the Mathematics Education Research Group of Australasia*, pp. 390-396.

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.

Paige, J. M. and Simon, H. A.: 1966, 'Cognitive processes in solving algebra word problems', in B. Kleinmuntz (ed.), *Problem solving: research, method, and theory*, Wiley, New York, pp. 51-148.

Robitaille, D. and Garden, R.: 1989, *The IEA study of mathematics II: Contexts and outcomes of school mathematics*, Pergamon Press, Oxford.

Stacey, K. and MacGregor, M.: 1994, 'Algebraic sums and products: Students' concepts and symbolism', in J. P. da Ponte & J. F. Matos (eds.), *Proceedings of the Eighteenth International Conference for the Psychology of Mathematics Education* (Vol. IV), University of Lisbon, Portugal: PME, pp. 289-296.

Stacey, K. and MacGregor, M.: (in press), 'Curriculum reform and approaches to algebra', in R. Sutherland (ed.), *Algebraic processes and structure*.

Sutherland, R.: 1991, 'Some unanswered research questions on the teaching and learning of algebra', *For the Learning of Mathematics*, **11**, 40-46.

Tall, D. and Thomas, M.: 1991, 'Encouraging versatile thinking in algebra using the computer', *Educational Studies in Mathematics*, **22**, 1-36.

Thomas, M.: 1994, 'A process-oriented preference in the writing of algebraic equations', in G. Bell, B. Wright, N. Leeson, and J. Geake (eds.), *Proceedings of the 17th Annual Conference of the Mathematics Education Research Group of Australasia*, pp. 599-606.

Ursini, S.: 1990, 'Generalization processes in elementary algebra: Interpretation and symbolization', in G. Booker, P. Cobb, and T. Mendicati (eds.), *Proceedings of the Fourteenth Conference, International Group for the Psychology of Mathematics Education*, pp.149-156.