MOLLIE MACGREGOR and KAYE STACEY

# STUDENTS' UNDERSTANDING OF ALGEBRAIC NOTATION: 11–15

ABSTRACT. Research studies have found that the majority of students up to age 15 seem 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 present evidence for specific origins of misinterpretation that have been overlooked in the literature, and which may or may not be associated with cognitive level. These origins are: intuitive assumptions and pragmatic reasoning about a new notation, analogies with familiar symbol systems, interference from new learning in mathematics, and the effects of misleading teaching materials. Recognition of these origins of misunderstanding is necessary for improving the teaching of algebra.

The *Concepts in Secondary Mathematics and Science* [CSMS] research project (Hart, 1981) provided evidence linking students' levels of understanding of algebraic letters to Piagetian stages of cognitive development and to IQ scores. It was concluded that most of the 13 to 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 misinterpret the notation in different ways and why they make certain errors. In this paper we take this next step. We show that some common misinterpretations can be explained by considering factors more accessible than cognitive level to diagnosis and possible remediation. 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.

*Educational Studies in Mathematics* **33:** 1–19, 1997.

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. When individuals were re-tested, parallel versions of the original items were used. 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, they are taught to use letters to stand for unknown or generalised numbers, frequently in the context of writing formulas for number patterns. They are given the 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 (see, for example, Assessment of Performance Unit, 1985; Booth, 1984; Cambridge Institute of Education, 1985; Herscovics, 1989; Küchemann, 1981; Robitaille and Garden, 1989).

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

1. The letter is ignored, given an arbitrary value, or used as the name of an object.
2. The letter is used as a specific unknown number or 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 IQ scores tended to demonstrate higher cognitive levels and made faster progress through algebra levels than students with lower IQ scores. Nevertheless the fact that a few students with below-average IQ scores reached the third or fourth algebra levels by age 15 (see Hart, 1981, p. 185, Figure 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 some of these other factors might be.

If a relatively unchangeable cognitive level is indeed a determinant of the way in which students can interpret letters and use algebra notation, then a successful introduction to algebra needs to take this fact into account. The pattern approach to algebra, currently presented in curriculum advice to schools (Stacey and 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 cognitive capacity of ordinary students. Additionally, capacities should not be considered outside a context for their expression. The success of students in experimental computer environments (Cohors-Fresenborg, 1993; Sutherland, 1991; Tall and Thomas, 1991) is evidence 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.

Küchemann (1981) concluded that the majority of 13 to 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 and to identify sources of difficulty.

## RESEARCH QUESTIONS

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 ways in which algebraic letters were interpreted by 11 to 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. Interviews with individual students at another school provided important insights into the causes of certain errors not previously reported in the literature. 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. We discuss two of the test items, shown in Figure 1, that 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 you write for Sam's height? .......... |

*Figure 1.* Pre- and post-test versions of Items 1 and 2.

*Expected responses*

We expected that in the pre-test most if not all 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 lower division (i.e., letter ignored, given a numerical value, or used as a label for an object). A possible lower-division 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 lower-division 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.

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, one of these students regarding alphabetical order as relevant (associating

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 | $h10$ | 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'. |

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 | $h10$ | $x$ | (Although 10 can be 'joined' to $h$, as 10$h$, 1 cannot be 'removed' from $y$. To denote 1 less than $y$, write $x$) |
| 1 | U$h$ | Uw | (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 1) |
| 1 | 'half' | [no response] | |

$g$ with $h$). 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 across both items. Table II shows the responses of these 14 individuals to the two items, and explanations for their responses to Item 2 (SUE).

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 lower division of interpretation. The last in the list cor-

TABLE III

Percentage of Year 7 students' responses to Item 1 in each of six categories before and after instruction

| Category | | Percentage of sample | |
|---|---|---|---|
| | | Pre-test ($n = 42$) | Post-test ($n = 38$) |
| *Unknown quantity* | [correct] | 2% | 37% |
| | [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% |

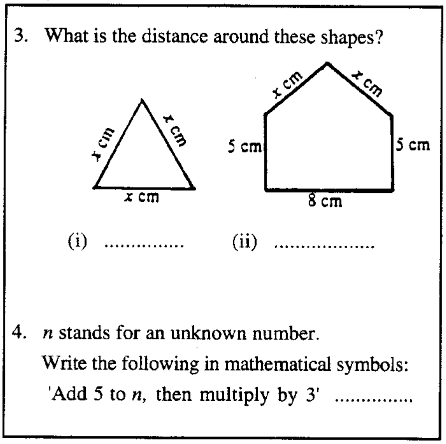
responds to Küchemann's higher division of interpretation (i.e., specific unknown or generalised number). Küchemann's hierarchy does not explicitly include the other two interpretations.

- *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 20 lessons. In this test, given to 38 of the original sample, 34 students responded and many were correct. As in the pre-test, almost all students reasoned consistently over both items.

As shown in Table III, one-third of the class (14 students) were correct, compared with only one student correct in the pre-test. For the Con's height item (see Figure 1) 10 students wrote the terms $8y$, $y8$, $8 - y$ or $y - 8$, in which they have made an attempt to denote a combination of the given number 8 and the unknown number $y$, their errors being due to conjoining terms for addition or writing the wrong operation. Combining these responses with the 14 correct ones, we see that 24 responses were not restricted to the lower division of interpretation. These 24 students have accepted that an algebraic letter does not stand for a word or need to be replaced by a number, and can be used along with a numeral to express an 'answer'. They understood the significance of algebraic letters


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' ..............

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

well enough to write an expression containing the symbols $y$ and 8. The percentage with this level of understanding had increased since the pre-test from 4% to 63%. The percentage of lower-division interpretations and omissions (29%) is much less than in the pre-test, and no student used a letter as an abbreviated word or person's name or assigned an alphabetical value. Only three students chose numerical values. Three chose other letters ($x$ or N), these being letters they had seen used for unknown numbers 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 (42%) and seven students (18%) using letters as unknowns but making syntax or operation errors (e.g., $x - 95$, $-x95$, or $x + 95$).

The performance of this class was better than expected, given the low success rates in published reports for similar items, and may indicate the effectiveness of the teaching program that had been used. As we will see in the next section, older and more experienced students often misinterpreted algebraic letters in other ways and many were not able to provide unclosed answers.

*Results from Large Sample of Students in Years 7 to 10*

The items DAVID and SUE that had been used for the small Year 7 sample were included in a test used by 22 schools for Years 7 to 10. All students in this large sample had been taught some algebra. We assume that students at all levels would have seen or written expressions coordinating two operations as required in Item 4 (referred to below as TWO OPS) in

TABLE IV
Percentages of students in Years 7–10 correct for Items 1–4 (N = 1463)

| Item [answer] | Year 7 ($n = 307$) | Year 8 ($n = 511$) | Year 9 ($n = 338$) | Year 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% |

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 Item 3 (referred to below as DISTANCE). The results for the four items are shown in Table IV.

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, even on the easiest item fewer than three-quarters of the Year-10 students were successful. Despite gains from one year to the next, the success rate for the hardest item (Item 4) did not reach 50%. 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.

Success rates on DAVID and SUE at Year 7 for this large representative sample were approximately the same as the rates for the previous small Year-7 sample on their post-test. Approximately half the Year-8 students and two-thirds of older students were correct for these very simple items. The six categories of interpretations previously identified were seen at all year levels. Three misinterpretations not seen in Year-7 responses appeared in the work of students in Years 8 to 10. They were:

- letter is a *label* associated with the name of an object (e.g., C to mean 'Con's height' and D to mean 'David's height' in $\mathrm{C} + 10 = \mathrm{D}$)
- *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 $h = h + 10$).

Furthermore, certain responses from older students showed evidence of interference from a variety of sources. These include new learning in mathematics and analogies with other symbol systems.

*Interference from new learning*

As we had observed in the first small Year 7 sample, younger students often ignored the algebraic letter and gave a numerical answer. For DAVID and SUE they chose an arbitrary number. For DISTANCE many of them measured the lengths marked '$x$ cm' with their rulers. They probably thought this was what the teacher wanted. On the other hand, older students gave numerical responses that were not due to these causes but reflected interference from new learning that had been misunderstood. Some errors were due to the *letter = 1* belief (explained below), and others were 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.

$$10 + h \qquad \frac{10 + h}{h} \times \frac{h}{h}$$

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

$$h = 5$$

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, numerical responses of this type have been misclassified as *letter ignored* or *arbitrary numeral*. Our analysis shows that complex (but incorrect) reasoning processes are likely to be involved.

Older students' belief that any letter stands for 1 was strongly evident in two of the 22 schools in our sample, and possibly present in others. For DISTANCE (ii), for example, we had seen the answer 20 cm in several written responses, in one instance with the added note 'because $x = 1$'. The prevalence of this belief was confirmed 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 $x$'. A student working on the DISTANCE item 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 test. Answers that we had first classified as arbitrary numerical value (e.g., David's height = 11) or attributed to inaccurate measurement (e.g., 20 cm for DISTANCE (ii)), could in many cases be now 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 source of confusion for older students is learning that the power of $x$ is 1 if no index is written and that $x$ with zero as the index equals 1 (i.e., $x = x^1$ and $x^0 = 1$). Like numerical responses caused by misguided attempts to simplify or solve (see above), numerical responses caused by the *letter equals* 1 belief may have been misclassified in previous studies as *letter ignored* or *arbitrary numeral*. We now know that at least some of them are due to the belief that a letter has the value 1. The concentration of this belief in particular schools points to a teaching effect.

Interference from new learning was also evident in the widespread misuse of exponential notation (e.g., $x^3$ instead of $3x$). This misunderstanding, seen in responses to DISTANCE as well as 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, supported by interview data, suggests that one major cause is lack of clear concepts for repeated addition, multiplication and repeated multiplication. (See Stacey and 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.

Other reasons why older students had more opportunities for making mistakes than younger ones was because of interference from new schemas only partly learned or because of their expectation of being able to use more advanced knowledge. When students were interviewed on the DISTANCE (ii) item, some of them made comments such as 'Do I have to figure out the numbers?' and 'That's the hypotenuse', and others began to draw lines in the figure to form two right-angled triangles. Their questions and methods indicated that they were searching for remembered schemas or learned procedures from geometry, and did not recognise that the task simply required a sum of measures. It was interesting to see that several Year 10 students wrote $x^2 + 5^2 + 8$. Their uncertainty about how to write 'twice $x$' (strongly evident in other items in the test) may have contaminated their knowledge of how to write 'twice five'. Alternatively, they may have tried to use knowledge about right-angled triangles and Pythagoras's theorem which they remembered concerns a sum of squares.

In TWO OPS, which required students to coordinate two operations using an unclosed expression, most students at all year levels were not correct (see Table V). The low success rate is in accordance with the findings of Ursini (1990) for a similar item used for students in Mexican schools, but considerably better than the success rates reported by Küchemann (1981, p. 108). In Table V the category 'Omit brackets' includes all responses that

TABLE V
Results for TWO OPERATIONS (N = 1806)

| Year | $n$ | Correct | Omit brackets | Use of exponent | $15n$ | $5n \times 3$ | 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% |

would have been correct if brackets had been inserted, for example $n+5\times 3$. According to Küchemann, success on this item requires interpreting the letter as a specific unknown or generalised number.

We had expected conjoining to be a common error, particularly for younger students. 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 or left a space for the 'answer' (e.g., $n+5 = x \times 3$ or $n+5 = \_ \times 3$). There were instances of exponential notation (e.g., $(n+5)^3$) and evidence of the 'unknown letter equals 1' belief which gives the answer 18 when the student adds 1 to 5 and then multiplies by 3. We have as yet no explanations for students' reasoning behind the many other forms of incorrect responses and the reluctance of so many students to write anything at all. We believe that the large percentage of unclassified responses and omissions points to students' lack of experience in using algebraic notation to express simple information or communicate their understanding.

*Analogies with other symbol systems*

The incidence of conjoining for addition (e.g. $10h$ to represent 10 plus $h$) was far lower than expected at all levels, being well below 10% after Year 7. However, a few of these 'conjoiners' appeared to believe that if a coefficient is of the left on the letter it indicates subtraction and if it is on the right it indicates addition. They wrote $h10$ to mean 'add 10 to $h$' and $1y$ to mean 'take 1 from $y$'. We assume that this notion may come from their experience with adding and subtracting along the number line or from their knowledge of the Roman numeration system in which VI means '1 more than 5' and IV means '1 less than 5', or that it may be based on deeper

TABLE VI

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

| Item | School A, Year 8–9 ($n = 70$) 1st | 2nd | 3rd | School B, Year 8–9 ($n = 60$) 1st | 2nd | 3rd | School C, Year 9–10 ($n = 26$) 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 |

intuitive metaphorical concepts associated with addition and subtraction (Lakoff and Johnson, 1980).

The alphabetical interpretation of letters, seen in data from several schools as well as in the first small sample of beginners, may come from prior experience with codes and puzzles, but it is interesting that this thinking also gave rise to 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 ($\alpha = 1, \beta = 2$, etc.). 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.

It is possible that students who gave the answer $h = h + 10$ had learned this symbolism from a computer language. There are many sign systems, formal and informal, from which students may draw analogies with algebra. These sources of interference or support have not yet been fully investigated.

### *Trends in Three Schools over 13 Months*

The effect of inappropriate teaching materials on students' interpretation of letters, particularly evident in certain schools, is discussed in this section.

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.

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.

*Interpretation of letters*

Misinterpretation of algebraic letters as abbreviated words or labels for objects was a persistent difficulty at School C over all year levels. 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 $5c$ 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 Schools A and B, 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.

The alphabetical interpretation of letters was also relatively common at School C. At Year 9, several students wrote R (i.e., ten letters after H) for DAVID and X (i.e., one letter before Y) for SUE. The teachers were unaware of this misunderstanding. When it was brought to their notice after the first test, it was easily corrected. In contrast, the ‘letter as abbreviated word’ belief was far more resilient and continued as a source of difficulty for students at this school.

*Coordinating two operations*

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 men and women’ the word ‘French’ modifies both ‘men’ and ‘women’. 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 sharply 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 and in other schools. 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.

## Discussion and Conclusions

We have shown that students frequently base their 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 and may persist for several years if not recognised and corrected. We suggest that younger students' misinterpretations 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 consolidate learning, a condition which has many causes.

The success of the Year 7 students in our sample indicates that many of them appear to have moved beyond or by-passed the lower-division interpretations. However, we do not know how far they had progressed along the route to understanding the significance of letters. At all year-levels there were some students who seemed to be unable to deal with the precise distinctions between letters 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). In everyday communication, such precise distinctions when referring to objects, attributes or measures are usually not important. However, wrongly interpreting an algebraic letter as the name of an object (e.g., interpreting $r$ to mean ‘red pencils’, so $6r$ 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 and 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.

There are many sign systems used in school and in everyday activities as diverse as electronics and knitting from which students may be drawing helpful and unhelpful analogies with algebra. The symbol system of chemistry, for example, resembles the notation of algebra in its use of brackets to refer to grouping and equations to represent the combining of quantities. Students have to learn the different meanings of conjoined letters in the two systems (addition in the sense of chemical combination in chemistry, but multiplication in algebra) and the different positions of numerals to indicate multiples of quantities. Some characteristics of sign systems are intuitive and easily learned, others are not. Even sign systems such as the way in which items are numbered in a textbook may influence students’ ideas.

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. The alphabetical interpretation was found to be common in one school where it had been reinforced by the use of puzzles and codes in mathematics lessons. The 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 this way. 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.

Our data show that after Year 7 a variety of misuse of algebraic notation arises as a result of interference from 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 algebraic 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.

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, the use of misleading teaching materials in early lessons, intended to make initial algebra learning easy, can seriously disadvantage students.

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 and write algebraic expressions. Attributing their errors to cognitive level alone accounts for failure to write correct answers but does not explain the wide variety of misinterpretations that we observed in our data. We have shown that some common misunderstandings are the results of particular teaching approaches, and can be avoided; others are developed by the students themselves. Certain misinterpretations can be reduced or overcome easily by appropriate teaching;

others are known to be very resilient. 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 in 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.

Lakoff, G. and Johnson, M.: 1980, *Metaphors We Live By*, University of Chicago Press, Chicago.

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 and J.F. Matos (eds.), *Proceedings of the Eighteenth Inter-*

*national 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* Mexico: PME, pp. 149–156.