Investigating Students' Numerical Misconceptions in Algebra

Vicki Steinle
University of Melbourne
<v.steinle@unimelb.edu.au>

Eugene Gvozdenko
University of Melbourne
<eugeneg@unimelb.edu.au>

Beth Price
University of Melbourne
<priceb@unimelb.edu.au>

Kaye Stacey
University of Melbourne
<k.stacey@unimelb.edu.au >

Robyn Pierce
University of Melbourne
<r.pierce@unimelb.edu.au>

Details are provided of simple algebraic items which can be used to detect two particular ways students think about the numerical meanings of letters (in contrast to non-numerical thinking). The data from Year 7 students (n=228) and Year 8 students (n=139) on these items is analysed using response patterns to probe student thinking. Less than 10% of the students were correct on these items whilst the prevalence of the two most common error patterns is 20%-30% at each year level. New response patterns are detected, indicating that further improvements to the items will enable students' thinking to be investigated further.

Many researchers have investigated students' difficulties with meaning of letters in algebra (for example, Küchemann 1981; MacGregor & Stacey 1997, Fujii 2003). For example, it is well-known that some students incorrectly interpret a letter as an object rather than a number, which we classify as *non-numerical* thinking. This is in contrast with incorrect *numerical* ways of thinking which involve replacing letters by numbers, but having incorrect ideas about what values the numbers can take.

In this paper, we provide additional evidence of the existence of such numerical misconceptions related to students' interpretations of letters. We show how patterns of responses to carefully designed items can reveal student's thinking.

Patterns of responses have been shown before, in another context, to reveal students' thinking. Steinle and Stacey (2003a, 2003b) showed that a specially designed Decimal Comparison Test provided data that was amenable to response pattern analysis and that large groups of students had misconceptions regarding the size of decimal numbers.

The items analysed in this paper come from a much larger set we are developing as an online resource (www.smartvic.com) to assist teachers with "assessment for learning". SMART tests (Specific Mathematics Assessments that Reveal Thinking) are designed for teachers to use for diagnostic purposes and planning future teaching (e.g. next day). These electronic tests are quick for students to complete (5 minutes) and the teacher is instantaneously provided with an online diagnosis for each student. Accompanying the diagnoses are explanations of the associated likely student thinking and/or reasons for the errors, along with teaching suggestions and links to other resources. These suggestions are designed to increase teachers' mathematical pedagogical content knowledge for specific topics, so affect both short-term teaching (of current students) and long-term teaching (of future students).

Literature

Küchemann (1981) is one of the foundational researchers in the field. Based on test responses of almost 1000 students (age 14) he analysed students' understanding of algebra; in particular, the various meanings that students ascribe to letters. He summarised students' apparent difficulties in understanding of the symbols or letters used in algebra at two main

In R. Hunter, B. Bicknell, & T. Burgess (Eds.), *Crossing divides: Proceedings of the 32nd annual conference of the Mathematics Education Research Group of Australasia* (Vol. 2). Palmerston North, NZ: MERGA.

levels. One of the items in this report by Küchemann, used later by Booth (1988), and Fujii (2003), asked students to decide when the statement L+M+N=L+P+N was true (always, never or sometimes). He found about half the students in his sample chose *never*; other researchers have likewise found a high proportion of students in their samples doing the same. This particular misconception is labelled *Different Letter means Different Number* (*DLDN*) in this paper.

MacGregor and Stacey (1997) conducted a large scale study of some hundreds of 11-15 year old students to identify students' interpretations of algebraic notation. They provide a list of reasons why students have difficulties with algebra:

> 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; and poorly-designed and misleading teaching materials. (p. 1)

Some of the interpretations that Küchemann (1981) and MacGregor and Stacey (1997) discuss involve non-numerical interpretations of letters, for example, *letter ignored* and *letter as object*. Other incorrect interpretations are *numerical*, for example, students who associate a number according to the position of letter in the alphabet (a = 1, etc), and those who believe that different letters must represent a different number (*DLDN*). In this paper we will focus on numerical misconceptions.

Such misconceptions are not confined to one culture nor to students who are generally 'weak' in mathematics. Fujii (2003) reports data from Japan – a country which tends to score well in international testing. He provided students with written test items including the two problems in Figure 1. Fujii noted that he had intended to analyse these two problems individually but, after he noticed groups of students who were correct on Problem 1 and incorrect on Problem 2 and vice-versa, he realised the importance of examining response patterns across both items.

Problem 1:	*Problem 2:*
Mary has the following problem to solve: "Find values(s) for x in the expression: x + x + x = 12" She answered in the following manner: a. 2,5,5 b. 10,1,1 c. 4,4,4 Which of her answer(s) is (are) correct? (Circle the letter(s) that are correct: a,b,c) State the reason for your selection.	Jon has the following problem to solve: "Find values(s) for x and y in the expression: x + y = 16" He answered in the following manner: a. 6,10 b. 9,7 c. 8,8 Which of his answer(s) is (are) correct? (Circle the letter(s) that are correct: a,b,c) State the reason for your selection.

Figure 1. Source items from Fujii (2003).

He allocated students to one of four groups based on their response patterns, as in Table 1. The first group (referred to as *Experts* in this paper, and Level 3 by Fujii) consisted of students who were correct on both items. He referred to another group as Level 1; these students circled all answers in both problems and hence were incorrect on Problem 1 and yet correct on Problem 2. We will label this misconception *Empty Boxes (EB)*, as these students appear to treat the letters x and y as they might treat empty boxes (i.e. interpret Problem 1 as □ + □ + □ = 12) in a primary school "fill the gaps" exercise. The next group (Level 2 by Fujii) correctly answered Problem 1 but were incorrect on Problem 2, by making a particular error; omitting to circle part c. We will label this misconception *Different Letter means Different Number (DLDN).* These students reject the option where x and y are equal. Response patterns which did not match those described were labelled as Level 0; Fujii suggested that these students "had a vague conception of literal symbols" (p.56).

In the last part of each problem, Fujii asked students to explain their reasoning. This written data provided evidence for the thinking behind the response patterns, as described above, which was further confirmed in interviews.

Table 1
Summary of response patterns to Problems 1 and 2 identified by Fujii (2003)

Problems	Level 3 (Expert)	Level 2 (DLDN)	Level 1 (EB)	Level 0 (other)
Problem 1	c√	c √	a,b,c ×	else
Problem 2	a,b,c√	a b ×	a,b,c √	

It is important to note that data analysis involving, for example, the percentage of students correct on individual items will not reveal the response patterns described above. Stacey and Steinle (2006) analysed data from another diagnostic test (DCT2) in two ways: using response patterns and Rasch analysis. They note,

> Being correct on an item for the wrong reason characterises DCT2. It is one of the reasons why the DCT2 data do not fit the Rasch model, because these items break with the normal assumption that correctness on an item indicates an advance in knowledge (or ability) that will not be 'lost' as the student further advances. (p. 87)

As can be seen from Table 1, there are students who can give correct answers to a problem for the "wrong reason"; for example, students with EB thinking can answer Problem 2 correctly. It is the pattern of responses, not the individual responses, that is the key for diagnosis.

Method

Participants and Procedure

The data reported in this paper comes from Year 7 and Year 8 students from four secondary schools in Melbourne, representing a range of socio-economic backgrounds. The testing took place in Term 2 of 2008; at this time, all Year 7 students from the four schools were tested and all Year 8 students at three schools.

While all students in a year level were involved in testing, only students who returned consent forms (57% overall) are included in this analysis. The normal classroom teacher took their complete class into a computer laboratory for the duration of a normal lesson for the purposes of the testing. These 2 items (with identifier codes 498 and 502) were offered amongst a large set of test items. Three different test versions were provided to reduce the opportunity for students to copy from those around them.

For ease of discussion, students who did not fully complete these two items are not reported below. Hence, the following analysis is based on 367 students who provided all 6 responses to the two items of interest: Year 7 (n=228) and Year 8 (n=139).

Research Instruments and Analysis Tools

The data discussed below is from two items within the Algebra Module; these items are provided in Figure 2 and were adapted from Fujii (2003). Firstly, we subtly changed the instructions. Rather than circling or selecting the option(s) which the students believed were correct, (this assumes that non-selection indicates the student thinks the statement is wrong), we provided three drop-down lists in each item, and students needed to actively choose each time.

Secondly, we changed the format of the last task in Item 498: *Mandy wrote x = 4* replaced the option *4,4,4* to make this statement more mathematically precise.

Item 498:	*Item 502:*
Some students had to find some values of x to make this equation true: $x + x + x = 12$ Mark the work of each student. Mary wrote $x = 2$, $x = 5$ and $x = 5$ Millie wrote $x = 9$, $x = 2$ and $x = 1$ Mandy wrote $x = 4$	Some students had to find some values of x and y to make this equation true: $x + y = 16$ Mark the work of each student. John wrote $x = 6$, $y = 10$ Jack wrote $x = 8$ and $y = 8$ James wrote $x = 9$ and $y = 7$

Figure 2. Items from SMART test Algebra Module

The response patterns that we predicted would be generated within the data are provided in Table 2. A Pattern Recognition script was created by the research team to assist with data analysis. This script identifies patterns in specified columns of the spreadsheet as well as the frequency with which each pattern occurs.

Table 2
Predicted response patterns and accuracy patterns to Items 498 and 502.

Abbreviated tasks within Items 498 and 502	Response Patterns			Accuracy Patterns		
	Expert[3]	DLDN[2]	EB[1]	Expert	DLDN	EB
Mary $x = 2, x = 5, x = 5$	W	W	R	√	√	×
Millie $x = 9, x = 2, x = 1$	W	W	R	√	√	×
Mandy $x = 4$	R	R	R	√	√	√
John $x = 6, y = 10$	R	R	R	√	√	√
Jack $x = 8, y = 8$	R	W	R	√	×	√
James $x = 9, y = 7$	R	R	R	√	√	√

Note: The Accuracy Patterns indicate when a response of R (Right) or W (Wrong) is actually correct (√).
(1, 2 and 3 refer to Fujii's Level 1, Level 2 and Level 3)

Results and Discussion

The percentage of students (n=367) who were correct on each task in Item 498 is 46%, 56% and 75%, respectively, and in Item 502, 86%, 59% and 75%, respectively. These statistics are provided for completeness and are not discussed further in this paper.

The Pattern Recognition script provided a list of the most common patterns ordered by frequency. The top seven patterns had frequencies of 74, 70, 29, 23, 17, 13 and 11; details of these patterns are provided in Table 3.

Table 3
The seven most frequent response patterns (n=367)

Abbreviated tasks within Items 498 and 502	Response Patterns						
	EB	DLDN	New 1	Expert	New 2	New 3	New 4
Mary $x = 2, x = 5, x = 5$	R	W	R	W	W	W	R
Millie $x = 9, x = 2, x = 1$	R	W	R	W	W	W	W
Mandy $x = 4$	R	R	W	R	R	R	R
John $x = 6, y = 10$	R	R	R	R	R	W	R
Jack $x = 8, y = 8$	R	W	R	R	W	R	R
James $x = 9, y = 7$	R	R	R	R	W	W	R
Frequency	74	70	29	23	17	13	11
Ratio*	12.9	12.2	5.1	4.0	3.0	2.3	1.9

* The ratio of observed frequency to expected frequency if all students chose randomly.

Only 23 students provided the predicted response pattern of Expert, making this the fourth most common pattern. It is clear from Table 3 that the two most frequently occurring patterns are those that match the predictions from the two main misconceptions *EB* (n=74), and *DLDN* (n=70). There are four new patterns which emerged from the data and will be discussed below.

The seven patterns noted in Table 3 account for 237 (65%) of the 367 responses. If students were randomly choosing between two responses for six tasks, then we would expect all 64 possible patterns to be well-represented. The ratio between the observed frequency and the expected frequency if all students were guessing at random is provided in the last row of Table 3; the first two response patterns (*EB* and *DLDN*) occur more than 12 times as often as would be expected if students' choice of answers was random.

The pattern labelled New 1 (n=29) is unexpected but more common than the Expert pattern. A comparison with *EB* reveals one different response. We propose that the change to the original item (part c in Problem 1 in Figure 1) has caused a split within the group of students who have the *EB* misconception. If students believe that we need any three numbers which add to 12 in Item 498, then $x=4$ (while correct from an Expert's point of view) is not a suitable answer, and these students would then choose Wrong (W) on this task. This is an example of how edits to an item to improve it from one point of view can result in unexpected reactions by students with other points of view.

The pattern labelled New 2 (n=17) is also unexpected. A comparison with *DLDN* reveals one different response; part c in Item 502. One possible explanation is that some students might believe that as the letter x comes before the letter y, then the number x must be smaller than the number y. This would lead to the response of W (Wrong) when presented with $x=9$ and $y=7$. This hypothesis can be tested in later versions of the items.

The pattern labelled New 3 (n=13) might be due to students who think that "all equations have one solution", hence the one solution for $x + x + x = 12$ is x is 4 and the one solution for the one equation $x + y = 16$ is x and y are 8.

The pattern labelled New 4 (n=11) could arise in various ways. A comparison with *EB* reveals one different response; likewise there is one different response to *Expert*. Hence, careless answering (i.e. inconsistent application of the student' own thinking) from two groups of students might be the reason that this pattern has emerged, but this is not definitive. Alternatively, there might be a currently unidentified misconception that generates this response pattern. As we currently have no direct evidence for these speculations, New 3 and New 4 will be grouped with the other "unclassified" response patterns with low frequencies in the following analysis.

Prevalence of Classifications by Year Level

On the basis of the above discussion it seems reasonable to group some patterns together for the next analysis. Four codes will be used; Expert is the code given to just one response pattern (as defined in Table 2) while the other three codes are given to various combinations. EB_total is a combination of EB and New 1; DNDL_total is a combination of DNDN and New 2 (making it a slightly larger group than Fujii would have identified); and Unclassified is a combination of all other response patterns including New 3 and New 4.

Figure 3 provides an analysis of the distribution of these four codes for Year 7 (n=228) and Year 8 (n=139). We can see that there is a small growth in the percentage of Experts (about 5% to about 10%) from Year 7 to 8. This surprisingly low rate is not dissimilar to those reported in Fujii (2003); about 10% of similarly-aged students in both the US and Japanese samples.

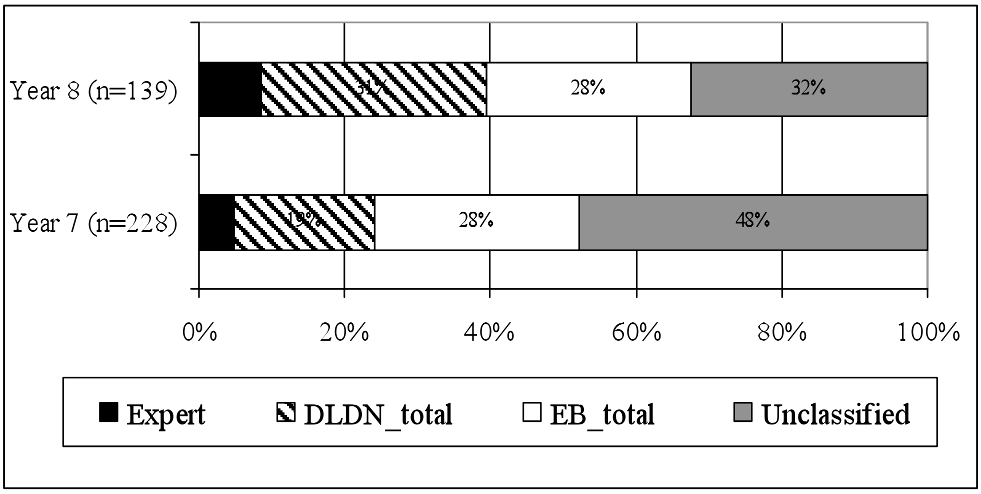

Figure 3. Percentage distribution of four codes in Years 7 and 8.

The percentage of students in EB_total is constant at about 30%, which is surprising given that the Year 8 students have had an extra school year of exposure to algebra. If this is, as Fuji predicts, a lower level of understanding than DLDN, we might expect a decrease in the prevalence of this misconception with older students.

There is an increase in the proportion of students in DLDN_total from about 20% (Year 7) to 30% (Year 8) which does accord with the suggestion that DLDN is a higher level misconception than EB. This is not longitudinal data, so this paper does not provide direct evidence.

There is a drop in the proportion of students in Unclassified from Year 7 to Year 8 (about 50% to about 30%). Fujii labelled those students Level 0, suggesting that they had a less sophisticated understanding than students with EB (his Level 1). We contend that the Unclassified group might contain students with varying levels of understanding. In our previous longitudinal research, we identified a hierarchy based on *readiness to move to* *expertise* (Steinle & Stacey 2003a). Of the students who were not Experts, those who were labelled Unclassified had a higher likelihood of moving to expertise on the next test (about 6 months later), than students who had an identified misconception. It seems that a student with a strongly held wrong belief is less likely to learn compared with another student who does not. This may also be true for understanding of letters.

Recall that our intention with these two items was to diagnose students with particular numerical misconceptions with the meaning of letters. What about students who have non-numerical interpretations (for example, think that *y* stands for *yachts*)? We expect that these students will be among the Unclassified group. If this is the case, then it would add weight to the hypothesis that these students have a less sophisticated understanding than EB or DLDN.

Implications

Making teachers aware of typical student thinking in a particular topic is likely to change their teaching. They might recognise the typical misconceptions in comments made by students and then target teaching to remove them, or, they might be able to prevent these misconceptions occurring in the first place. Research on changes to teachers' mathematical pedagogical content knowledge is being conducted in this project.

Clearly there is further research which can be done to resolve some of the issues raised in this paper. For example, when we analyse the other items in the SMART test's Algebra Module that these students completed we can determine (i) if, based on responses on other items, there is a hierarchy in these misconceptions, (ii) what proportion of the Unclassified students hold non-numerical misconceptions, (iii) how the students labelled as Experts on these items deal with other algebra items.

Furthermore, in the light of these results, we have revised our items for the next round of testing. We have included additional statements for students which will allow us to identify the causes of New 1 and New 2 patterns.

Conclusion

> It is only by asking the right, probing questions that we discover deep misconceptions, and only by knowing which misconceptions are likely do we know which questions are worth asking (Swan, 1983, p 65)

The work presented in this paper has attempted to break into this cycle by further investigation of patterns of student thinking identified by earlier research (Küchemann 1981; MacGregor & Stacey 1997, Fuji 2003). In addition, we have adapted items from a paper-test to an online test that allows for editing or addition of items in later testing cycles. The analysis of the data using response patterns has confirmed predicted misconceptions about the numerical values that letters may be assigned, and has revealed some interesting variations. These new hypotheses will be investigated in future SMART testing.

Response pattern analysis has already proved useful in diagnosing student misconceptions in students' understanding of decimal notation (Steinle & Stacey 2003a, 2003b). In this paper we have shown how another topic has been amenable to the same student-focussed approach, in contrast with other approaches such as item analysis or student scores.

The prevalence of expertise on these simple algebraic items is very low; about 10% of the Year 8 students chose correctly on the set of six tasks. For approximately 90% of these

Year 8 students, then, mathematics lessons containing algebra are rendered incomprehensible; these students are trying to learn procedures, without meaning, carried out on letters with the wrong meaning. For example, a student with the *Empty Box* misconception, who interprets $x + x + x = 12$ as $\square + \square + \square = 12$ and believes that 10+1+1 is an acceptable answer, is unlikely to make sense of the explanation that $x + x + x= 12$ is equivalent to $3x=12$ and then $x=4$. The consequences of limited conceptual understanding of the meaning of letters are quite severe. The possibility of using technology to support someone with limited procedural skills is unlikely to work if they have inadequate conceptual understanding. Computer algebra systems still require an understanding of the meaning of letters and the correct interpretation of variables.

Acknowledgements

The items and data reported in this paper come from an Australian Research Council - funded project with the Department of Education and Early Childhood Development (Victoria).

References

Booth, L. (1988). Children's difficulties in beginning algebra. In A. Coxford & A. Shulte (Eds.) The ideas of algebra, K-12. 1988 NCTM yearbook. Reston, VS: NCTM.

Fujii, T. (2003). Probing students' understanding of variables through cognitive conflict problems: Is the concept of a variable so difficult for students to understand? In N. Pateman, G. Dougherty, J. Zilliox (Eds.), *Proceedings of the 27th annual conference of the International Group for the Psychology of Mathematics Education,* (Vol.1, pp. 49-65). Hawaii: PME.

Küchemann, D. (1981). Algebra. In K. M. Hart, M. L. Brown, D. E. Kuchemann, D. Kerslake, G. Ruddock, & M. McCartney (Eds.), Children's Understanding of Mathematics :11-16 (pp. 102-119). Oxford, U.K.: John Murray.

MacGregor, M. & Stacey, K. (1997) Students' understanding of algebraic notation: 11-16 *Educational Studies in Mathematics.* 33(1), 1-19.

Stacey, K. & Steinle, V. (2006). A case of the inapplicability of the Rasch Model to mapping conceptual learning. Research in Mathematics Education: Insights from Rasch Measurement, A special issue of the Mathematics Education Research Journal, 18(2), 77-92.

Steinle, V. & Stacey, K. (2003a). Grade-related trends in the prevalence and persistence of decimal misconceptions. In N.A. Pateman, B.J. Dougherty & J. Zilliox (Eds.), Proceedings of the 27th Conference of the International Group for the Psychology of Mathematics Education, 4, 259 - 266. Honolulu: PME.

Steinle, V. & Stacey, K. (2003b). Exploring the right, probing questions to uncover decimal misconceptions. In L. Bragg, C. Campbell, G. Herbert & J. Mousley (Eds.) Mathematics Education Research: Innovation, Networking, Opportunity. *Proceedings of the 26th Annual Conference of the Mathematics Education Research Group of Australasia*, 2, 634-641. Sydney: MERGA.

Swan, M. (1983). Teaching decimal place value: A comparative study of conflict and positive only approaches. Nottingham: Shell Centre for Mathematical Education.