How Important is the Reversal Error in Algebra?

Francis Lopez-Real

The University of Hong Kong

Abstract

A sample of 577 secondary pupils in Hong Kong were tested on a set of problems that commonly give rise to the 'reversal' error. Two areas of the test are discussed in this paper: (i) the effect of the syntactic structure of a question with respect to the contiguity of elements in the sentence, and (ii) the effect of the inclusion of subsidiary questions within a problem. The results are discussed with particular reference to the construction of cognitive models based on comparison.

Introduction

The specific algebraic error known as the 'reversal error' has received a great deal of attention in research studies over a period of more than ten years. There are two broad categories of reversal error which may be described as the multiplicative and additive models. For example, 's is 5 times t' coded as '5s=t' and 's is 5 more than t' coded as 's+5=t'. The main reason for this attention has been its apparent resilience when attempts have been made to isolate and remove hypothesised causes. Explanations have been suggested in terms of the misapplication of natural language rules (Kaput, 1987), and direct syntactic translation (Mestre, 1988). Davies (1984) proposed an explanation in terms of 'frame retrieval' whereby the error is caused by selecting the 'label or unit' frame in favour of the appropriate 'numerical variables' frame. This explanation has a certain intellectual appeal. However, analysis of a typical algebraic curriculum would probably reveal examples of the 'label or unit' frame to be the *least* frequently encountered. Nevertheless, in the author's experience, the imagistic power of introducing algebra to children via the labelling model still seems to hold great appeal for many teachers. (In Hong Kong, for example, the most frequently used text-books introduce algebra through a numerical variables approach and yet many teachers still augment this with the typical 'a' for apples description when it comes to collecting like terms.) However, as McGregor (1991) has pointed out, the frame retrieval model does little to explain why the reversal error is still persistent in examples like 'y is 8 times z'. Clement (1982) explained the Student-Professor reversal ('There are 6 times as many students as professors' being written as 6S=P) as an association of six students per professor. This was described as 'static comparison' and the idea has been further developed into a convincing theory of cognitive models by McGregor & Stacey (1993) which also attempts to explain the occurrence of the reversal error when concrete referents are not used : 'These reversals appear to be direct representations of cognitive models in which the numeral is associated with the larger variable' (p.228).

However, another possible factor influencing the occurrence of the reversal error in such cases is what may be described as *contiguity*. This is not to suggest a simplistic word-by-word syntactic processing but rather a combination of the semantic and syntactic breakdown of a sentence into meaningful 'chunks' or 'sentoids' (Aitchison, 1989). Now the grammatical breakdown of the sentence 'p is 6 more than q' is the Noun-phrase, Verb, Noun-Phrase structure. That is, 'p' - 'is' - '6 more than q'. However, it is possible here that 'p is 6 more' may be processed as the first meaningful 'chunk' of the sentence. In natural language 'is' may frequently be synonymous with 'becomes' and thus, although 'p is 6 more' does not logically stand alone, 'p becomes 6 more' *can* be meaningfully interpreted and its symbolic representation would be p+6. In this event, would the re-arrangement of the sentence so that the '6 more' is separated from p have an effect on student performance? This is one of the questions addressed in the present study.

Another aspect that has frequently puzzled the author is the following. When discussing research findings on errors in algebra with serving teachers, a common response has been the confident assertion that their own students would not perform as badly as the quoted figures. Why is this? One's first reaction is that this is simply a case of wishful thinking and self-deception. After all, there are many examples of classroom studies reporting the surprise expressed by teachers when confronted with protocols of their own behaviour. Perhaps this is a similar phenonemon. But another reason can be hypothesised. In the classroom situation problems are rarely given to pupils in the isolated context that is typical of many research studies. A common structure for problems is the build-up to the required final result through a series of subsidiary questions. What difference to performance would such a structure have in the context of 'reversal-type' problems? This is another question addressed in the study.

Comparative Test Items

Three 'parallel' tests were constructed, each consisting of a total of ten questions. (The tests are designated below as A, B, and C). This paper analyses and discusses those items pertinent to the questions raised in the preceding section. Two sets of questions (additive and multiplicative) were used to test contiguity. These are described as Non-referent items since no concrete objects are referred to. An item matching one used by MacGregor & Stacey (1993) and using the phrase 'is the sum of' was also included in these sets because its high facility in their study was not satisfactorily explained by the 'cognitive model' theory. A similar item was also included for the multiplicative case. The matching items for this section of the test are given below (each statement was followed by the sentence: 'Write an equation that describes the above sentence'). The C items are certainly a little awkward linguistically but there seemed no other way to test the contiguity factor.

A1. p and q are numbers. p is the sum of 6 and q.

B1. p and q are numbers. p is 6 more than q.

C1. p and q are numbers. 6 more than q is the same as p.

A2. s and t are numbers. s is the product of 4 and t.

B2. s and t are numbers. s is 4 times t.

C2. s and t are numbers. 4 times t is the same as s.

There were a total of four items investigating the aspect of 'lead-up' or 'subsidiary' questions. These are described as Concrete-referent questions. The four items can be classified as Multiplicative/Descriptive; Multiplicative/Tabular; Additive/Descriptive; and Additive/Tabular. In the A test each item involved two subsidiary questions, the first purely numerical and the second involving just one variable. The B test used just the single-variable subsidiary question, and in the C test there were no subsidiary questions. It is therefore only necessary here to list the items from the A test.

A7. In a college there are 10 times as many students as teachers.

If there are 7 teachers, how many students are there? {Numerical}

If there are N teachers, how many students are there? {Single-variable)

If there are N teachers and M students,

write down an equation showing the relation between N and M. {Two-variable)

A8. The following table shows the cost of meat in the market:

Weight of meat (in Kg)	1	2	3	4	5
Cost of meat (in $)	20	40	60	80	100

What will be the cost of 8 kg of meat?

What will be the cost of P kg of meat?

If the cost of P kg of meat is Q dollars,

write down an equation showing the relation between P and Q.

A9. In a classroom there are 6 more girls than boys.

If there are 15 boys, how many girls are there?

If there are N boys, how many girls are there?

If there are N boys and M girls,

write down an equation showing the relation between N and M.

A10. The following table shows the weight of a suitcase packed with different amounts of clothing:

Weight of clothing (in Kg)	8	10	12	14
Weight of *full* suitcase (in Kg)	13	15	17	19

If the weight of the clothing is 17kg, what will be the weight of the *full* suitcase?

If the weight of the clothing is Q kg, what will be the weight of the *full* suitcase?

If the weight of the clothing is Q kg and the weight of the *full* suitcase is P kg,

write down an equation showing the relation between Q and P.

The Sample

A total of 577 Form 2 pupils from 6 secondary schools in Hong Kong were tested for the study. Hong Kong has a selective system at the transition stage from Primary to Secondary schools and pupils are assigned to one of 5 band-levels according to their academic performance at the end of primary education. Secondary schools will then cater for a narrow range within this band structure (e.g. a school may be described as a Band 2 & 3 school etc). Schools may opt to use English or Cantonese as the medium of instruction. At the present time about 70% of the secondary schools are *described* as operating an English-medium policy although the move appears to be in the direction of more Cantonese-medium schools in the future. (There is a strong perception, particularly among parents, that English-medium carries more prestige and is likely to be more useful for further academic studies.) In practice, however, very few schools adopt a 100% English-medium policy. In the case of mathematics it is very common for teachers to use a mixed-code presentation whereby most of the explanations are given in Cantonese with key phrases, especially those involving technical terms, emphasised in English. Nevertheless, the pupils in such schools use English language textbooks and take their examinations and tests in English. The schools chosen for this study were English-medium schools covering Bands 1 to 3 of the ability range. Hence the pupils were 'average to above-average' in terms of academic performance. For each class used in the study the three parallel tests were randomly distributed so that about one-third of the class answered each test.

Results

(i) Non-referent Items

The following table gives the results for the parallel questions of the Non-referent items.

Question	Number Correct	Total	Facility (%)	Errors	Reversal Error	Reversal as % of Errors
A1	176	192	92	16	3	19
B1	148	191	77	43	21	49
C1	167	194	86	27	3	11
A2	139	192	72	53	2	4
B2	155	191	81	36	21	58
C2	186	194	96	8	0	0

Given the high facility rates for all these questions, some care needs to be exercised in the interpretation of the results. Nevertheless, it is the *comparison* between parallel questions that is pertinent here and some interesting features do emerge. First, although A1 had the highest facility rate of its three parallel items, the structurally identical question for the multiplicative case produced exactly the reverse result. Why was this? Second, the comparison of the items testing contiguity show a statistically significant difference at the 5% level for the additive items (B1,C1) and at the 1% level for the multiplicative items (B2,C2). These results are discussed further in the next section.

(ii) Concrete-referent Items

The table below shows the facility rates for the four items covered in this section. The results for the subsidiary questions are not shown here since their purpose was to test their effect on the outcome of the *final* answers.

Question	Number Correct	Total	Facility (%)	Question	Number Correct	Total	Facility (%)
A7	126	192	66	A8	130	192	68
B7	137	191	72	B8	145	191	76
C7	97	194	50	C8	99	194	51
A9	147	192	77	A10	133	192	69
B9	152	191	80	B10	131	191	69
C9	116	194	60	C10	110	194	57

Although the facility rates vary, the *pattern* of results is consistent across multiplicative/additive items and descriptive/tabular. Curiously, the B items (with only one subsidiary question) had the highest (or equal highest) facilities throughout. Nevertheless, the differences between the A and B paired items were not found to be statistically significant. We can conclude that the inclusion of the numerical subsidiary question makes no *additional* contribution to the single-variable subsidiary question and hence, in order to see clearly the comparison between items *with* subsidiary questions and those *without*, in the following table the results for the A and B tests are combined.

The differences between all four paired items are statistically significant at the 1% level and, except for item 10, the increase in facility levels with the inclusion of subsidiary questions is around 20%. In common with similar studies the reversal error was by far the most frequent error pattern in items 7,8 and 9. However, the occurrence of the reversal error expressed as a percentage of all errors was dramatically reduced in the items with subsidiary questions compared to those without. Again these differences (for the paired items 7,8 and 9) were significant at the 1% level. In these three cases the reduction in the relative occurrence of the reversal error was over 20%.

Question	Number Correct	Total	Facility (%)	Errors	Reversal Error	Reversal as % of Errors
AB7	263	383	69	120	53	44
C7	97	194	50	97	63	65
AB8	275	383	72	108	33	31
C8	99	194	51	95	50	53
AB9	299	383	78	84	24	29
C9	116	194	60	78	44	56
AB10	264	383	69	119	14	12
C10	110	194	57	84	12	14

Discussion of Results

Returning to the Non-referent items, we must first consider why the item using the phrase 's is the product of 4 and t' was so much less successful than its additive equivalent. The most likely explanation is that the word 'product' was simply unfamiliar to many of the students. Whilst all the teachers involved had claimed that the word had been met and was used in the textbooks it still seems likely that its use would be far less frequent than the word 'sum' and hence, particularly for second-language students, a source of difficulty. The fact that the most frequent error in this case was the answer '4+t' or its equivalent, lends support to this explanation.

Taking the parallel items for question 1 together, can the results be explained in terms of cognitive models? There is no conflict with such a theory as long as we assume that different syntactic structures are likely to *produce* different cognitive models. Notice that this is *not* the same as saying that syntactic translation of sentences is automatically applied. By its syntactic structure, B1 produces a 'comparison relationship' cognitive model which is apparently more difficult for students to translate into mathematical terms, whereas A1 produces an 'equality relationship' model that is easier to process. For B1, as suggested in the introduction, in trying to translate into mathematical terms children may interpret the phrase 'p is 6 more' as a meaningful 'chunk' which leads to the reversal error. The re-arrangement of the phrase in C1, so that the 6 is no longer contiguous with the p, has not only raised the facility level but, perhaps more importantly, has significantly reduced the relative frequency of the reversal error. (This is even more dramatic for B2 and C2 where the reversal error disappeared entirely). Why is this? In line with the previous explanation, it would appear that the re-structuring of the question now leads to the equality cognitive model rather than the comparison model.

What are the implications? The language descriptions in items B1 and B2 are in fact much more likely to be encountered than the rather more awkward phrasing of C1 and C2 and we must help children cope with such situations. What is necessary is admirably expressed in the following quote from MacGregor & Stacey (1993):

> *'Students should be made aware that some relationships .... are easy to express in natural language and easy to comprehend, but must be paraphrased, reorganised, or reinterpreted before they can be expressed mathematically' (p.229).*

Turning now to the Concrete-referent items, one's first reaction might be that since the *purpose* of subsidiary questions is to help the student towards the final solution then the results are only to be expected and are of no interest. However, certain issues are inevitably raised by the results. First, if such a dramatic improvement can be effected by this device then just how serious *is* this type of error? (More of this in a moment).

Second, given that the subsidiary questions are likely to improve overall facility levels, this does not explain why the *relative* frequencies of the reversal error are so greatly reduced. Just *how* do the subsidiary questions help to achieve this? Taking item 7 as illustration, one might expect that the response to the second subsidiary question (i.e. 10N) would, in the cases where the final answer is *correctly* given, lead to that answer being expressed as M=10N. However, there were many cases (13% for AB7) where a 'Non-standard' correct answer was given. Examples are: M/10 = N; M/N = 10; M:N = 10:1; and even M-N = 9N (!). This suggests, at least for these incidents, that the function of the subsidiary questions may be rather similar to the ideas in the quote above i.e. to slow the pupils down so that they think about the *meaning* of the question and perhaps re-organise the statements of the question. Of course, there is no way of knowing that this does not also occur with pupils giving the 'standard' correct answer. The percentages of Non-standard correct answers for the other questions are: AB8 (14%); AB9 (19%) and AB10 (10%). As mentioned before, the case of item 10 was a little different to the other three since here the most frequent error was *not* the reversal error but variants on the response P=8Q/13. This suggests that the tabular form led many pupils to think that the relationship must be multiplicative, perhaps because that type has been most frequently encountered before. Having assumed this, their answer is based on some combination of the first numbers in the data.

There is one curious feature in the results that should be commented on here. Although the A,B items have been compressed because there was no significant difference in their *overall* results, nevertheless there *were* interesting differences in their error patterns and frequencies of non-standard correct responses as shown in the table below. It appears that the B test (which went directly to the single-variable subsidiary question and omitted the numerical) not only produced a far lower frequency of reversal errors but also produced generally higher frequencies of non-standard correct answers. It is difficult to understand why this should be so and perhaps needs further investigation.

Question	Total Errors	Reversal Errors	Percent	Total Correct	Non-standard Correct	Percent
A7	66	33	50	126	17	13
B7	54	20	37	137	18	13
A8	62	24	39	130	16	12
B8	46	9	20	145	22	15
A9	45	18	40	147	23	16
B9	39	6	15	152	35	23
A10	59	11	19	133	8	6
B10	60	3	5	131	19	15

Of course, an important objective in teaching is to enable the students to become *autonomous* users of mathematics, independent of the type of helpful props given in this study. The results here suggest that the cognitive models constructed by students are a function of the syntactic structure and lexical items of a statement as well as its semantic content. This in turn suggests that one emphasis in teaching should be on encouraging students to ask *themselves* subsidiary questions (similar to the sub-goal strategies useful in problem-solving) and to re-phrase statements in their own words.

Wider Issues

There are a number of problems with this kind of research, the most obvious being that while the results can tell us something about general *patterns* of response nothing can be inferred about *individual* reasons for a response. There is *always* a reason for any given pupil's answer, including putting down the first thing that comes into their head because they feel no commitment or have no vested interest in the outcome. This is particularly the case in large sample testing where anonymity is guaranteed and the pupils know there is no comeback, whatever their performance. Another problem is the 'snapshot' effect. That is, the type of questions posed often give no opportunity for further reflection or adjustment. How many professional mathematics educators would care to have judgements made about their own *initial* attempts at solving a problem, say? This may be especially pertinent with the kind of algebraic problems being tested here. Consider item 7 on the test, for example. What is the *purpose* in writing a relation between N and M? It certainly cannot be to enable one to calculate M for a given value of N, since this is easily calculated from the initial statement. In fact, the facility rate for the numerical subsidiary question of this item was 97% ! Writing the algebraic relationship between the two variables only really makes sense when it is part of a wider and more complex problem. In other words, it is likely to be a sub-goal of a problem. But here the question is divorced and isolated from any such framework and this is true of much research into algebraic problems. In linguistics, some studies on making sense of ambiguous sentences have suggested that a back-tracking and checking process is constantly being employed (Aitchison, 1989). It may be that similar processes need to be consciously employed when transforming natural language into mathematical form. Given the reservations above, the results of the present and other similar studies can still point us in the right direction as far as teaching is concerned. The most important factors appear to be the need to re-formulate, re-phrase, re-interpret, adjust and check initial attempts at transforming ordinary language into mathematical language. In a phrase, to become *reflective students* of mathematics. And hence, as far as research is concerned, perhaps further work in this area needs to concentrate *not* on what students cannot do but rather on what they *can* do in terms of self-correction and reflection.

References

Aitchison, J. (1989). *The Articulate Mammal.* London: Routledge.

Clement, J. (1982). Algebra Word Problem Solutions: Thought Processes Underlying a Common Misconception. *Journal for Research in Mathematics Education, 13*(1), 16-30.

Davis, R. (1984). *Learning Mathematics. The Cognitive Science Approach to Mathematics Education.*

London: Routledge.

Kaput, J. (1987). Towards a Theory of Symbol Use in Mathematics. In Janvier, C. (Ed.), *Problems of Representation in the Teaching and Learning of Mathematics.* Hillsdale: Erlbaum.

MacGregor, M. (1991). *Making Sense of Algebra: Cognitive Processes Influencing Comprehension.* Deakin University

MacGregor, M. & Stacey, K. (1993). Cognitive Models Underlying Students' Formulation of Simple Linear Equations. *Journal for Research in Mathematics Education, 24*(3), 217-232.

Mestre, J. (1988). The Role of Language Comprehension in Mathematics and Problem Solving.

In Cocking, R. & Mestre, J. (Eds.), *Linguistic and Cultural Influences on Learning Mathematics.* Hillsdale: Erlbaum.