Algebra, Symbols, and Translation of Meaning

Lillie Crowley
Engineering Division
Lexington Community College
Kentucky, USA

Michael Thomas
Department of Education
University of Auckland
New Zealand

David Tall
Mathematics Education Research Centre
University of Warwick
United Kingdom

In translating problems from a verbal description to algebraic notation, we consider what children will do with statements that do not have the syntactic constructions that may provoke the reversal errors of the "students–professors problem." MacGregor & Stacey (1993) show that children do not follow a simple word-matching order and postulate the existence of mental models that can be accessed in any order. We hypothesise that the children's responses are a natural consequence of their previous development and show that, as the verbal problems become more complex, children are more likely to revert to a process-orientation with the arithmetic operation on the left, rather than an assignment order with the variable on the left. We place this in a theory of cognitive compression as the children grow in mathematical sophistication.

Introduction

The "students and professors" problem of Clement, Lochhead and Monk (1981) has unleashed a profusion of articles formulating and testing theories about conceptualisation of symbols in algebra (see Laborde 1990 for a summary). Given the problem of translating "There are six times as many students as professors" into symbols using S for the number of students and P for the number of professors, 37% of college students sampled were in error and two thirds of these wrote $6S = P$, rather than $S = 6P$. This *reversal error*, has been identified by some as occurring because the letters are thought to stand for the objects (students, professors) rather than the *number* of objects and Kaput (1987) suggested that this was strongly influenced by the underlying language *syntax* where the "$6S$" suggested an adjective-noun structure meaning "six students". This leads to a *word order matching error* (or *syntactic translation*) which essentially follows the word order of the original problem (e.g. Schoenfeld, 1985; Mestre, 1988). Herscovics (1989) contrasted syntactic translation with *semantic translation*, which attempts to interpret the underlying meaning.

MacGregor & Stacey (1993) asked students to translate more simple statements into algebra and reported that there were responses, including errors which could not be adequately described by syntactic translation. They concluded that the students must have "cognitive models" including some easily described in natural language (such as "I have $6 more than you") but less amenable to translation into mathematical code:

> The variety and form of students' responses leads us to infer some properties of their cognitive models and to postulate that information from these cognitive models can be retrieved in any order. Such a retrieval process would explain the apparently random choice of responses that match or do not match the word order. (MacGregor & Stacey, 1993, p. 228.)

The evidence that we have partly supports this view. However, we do not consider that *all* children have cognitive models which can be retrieved in any order. Instead we suggest that the models developed by children are a result of their previous experience which requires several reconstructions to allow them to develop models with a flexible use of order of access. This relates to their learning experiences in mathematics, their use of symbols, and the natural process of compression of knowledge with growing maturity.

A theory of learning in arithmetic and algebra

We begin with the simple observation that the human brain has a huge long-term memory, but a limited short-term working store. It can therefore store vast quantities of information, but can only manipulate a small number of items consciously at any time. To cope with this limitation in mathematics, two strategies are adopted, one to routinise procedures so that they can be performed using little conscious memory, and a more powerful one using symbols to label complex information allowing a small number of symbols to be the focus of attention at any one time. These symbols are used by experts in a particularly flexible way in arithmetic and algebra by labelling a *process* with a *symbol* which then is also used for the *concept* produced by that process.

Perceiving a symbol as either process or concept gives great power to the individual, for the *process* enables him or her to *do* mathematics, but the *concept* allows him or her to *think* about it and manipulate it mentally. A symbol standing for both process and the output of that process is called a *procept*, with the additional factor that two different symbols which represent the same object can be regarded as the same procept (Gray & Tall, 1994). Mental manipulation of procepts gives the thinker great power. The flexibility in thought grows over time as the individual compresses the process to allow it to be thought of as an object. The proceptual child grows to regard different symbols for the same thing as being essentially the same object, whilst the procedural child persists longer in regarding them as different processes (Gray 1993).

Early arithmetic involves the *process of counting* which becomes compressed into the *concept of number* with *number symbols* fulfilling the pivotal role. The expression “5+4” can mean different things to different children at different stages, including:

(a) count-all (count 5 objects, then 4, then count them all),
(b) count-on (count on four starting after 5: “six, seven, eight, *nine*”),
(c) known fact (the answer is 9),
(d) derived fact (e.g. “I know 5 and 5 is 10, so 5 and 4 is one less”).

A child may have one or more of these interpretations at a given time. Those who only count-all may know a few facts, but are unlikely to derive facts (Gray & Tall, 1994). Some children remain procedural and prefer the security of counting-on to solve numerical problems. More successful children leave count-all behind and compress knowledge with great flexibility and fluency using a combination of (b), (c) and (d).

Depending on the development of the child in this sequence of learning, the following equations may have very different meanings:

(i) 5 + 4 = □, (ii) 5 + □ = 9, (iii) □ + 4 = 9, (iv) □ = 5 + 4.

The first of these can be answered by any of the four named processes of addition, the second is difficult with count-all, but straightforward with the other techniques (for instance, count-on from 5, and count how many are required to reach 9.) The third is difficult with count-on because the child does not know where to start the count (Foster in prep.). The fourth may make little sense to any child who reads an equation as a left-to-right process to given an answer (Kieran, 1981). For such children there is a preference to express the equation 9=5+4 as an addition process 5+4=9 (5 and 4 *makes* 9). In practice, the difference between (ii) and (iii) is ephemeral, and children soon see it to be equivalent to a subtraction, but (iv) continues to surface strongly later in algebra.

Many children have difficulty with a symbol such as "*x*+3" which they will not accept as an answer because they expect a number (Kuchemann, 1981). From our viewpoint, such children see the symbol *x*+3 as a process and not a mental object – a process they cannot carry out because they do not know what *x* is. To be able to cope with such a symbol requires not only that it be given a meaning, but that the meaning should cope with it both as a process (of evaluation when *x* is known) and also as an object which can be manipulated as it stands. It requires the flexibility which views the symbol as a procept.

Such a view is not available to a child who regards number operations only as counting procedures. Regrettably, in traditional teaching, faced with children in difficulty and lacking in comprehension, the way out is often "fruit salad" algebra. The symbol "3*a*+4*b*" is explained to stand for "three apples and four bananas". Some children who play along with this delusion are able to sort out "3*a*+4*b*+2*a*" as "three apples and four bananas and two apples", which is "five apples and four bananas", or "5*a*+4*b*". So they appear to be able to simplify expressions. But they now have an image of a letter as representing an object, set up ready to fall foul of the student-professor problem. Other children might simply conclude that "3*a*+4*b*+2*a*" is "nine apples and bananas" and – since they have no mathematical symbol for "and" – they may write the letters one after another, *conjoining* them as "9 *a b*". (see, for example, Booth, 1984).

Algebra teaching further exacerbates the differences between proceptual thinkers with their flexible use of symbolism and procedural thinkers who try to give it a temporal, process meaning. The equation symbolism is reversed to express *y* as a function of *x* in the form *y*=*x*+4. This *assignment order* causes no problems to the proceptual thinker but violates the meaning of the "equals" sign for the procedural child who may prefer the *process-oriented order* *x*+4=*y*. Aesthetic preferences are introduced – for instance, it is "usual" to write *y*=3*x*+4 rather than *y*=4+3*x*, although it is preferable to write *y*=4–3*x* rather than *y*=–3*x*+4, to avoid starting an expression with a minus sign. Flexible thinkers are more likely to take such things in their stride, coping with both the flexibility of meaning of the symbols and arbitrary matters of taste.

We hypothesise that proceptual thinkers will have cognitive models that show flexibility in meaningful re-ordering of algebraic symbols, coloured by personal preferences based on aesthetics and personal experiences. But procedural thinkers are more likely to prefer a *process-oriented order* which we hypothesise will become more dominant as the complexity of problems increases.

Empirical Data

The first problem of MacGregor and Stacey (1993) is:

> z is equal to the sum of 3 and y. Write this information in mathematical symbols.

The responses from 255 Year 9 students (aged 13 to 14) of mixed ability in Australia and their interpretation of the responses as (possibly) syntactic or not are as follows.

Response (possibly syntactic)	Frequency
(a) $z = 3 + y$	70
(c) $z = y + 3$	0
(e) $z = 3y$	37
(g) [unclassified]	27

Response (not syntactic)	Frequency
(b) $3 + y = z$	66
(d) $y + 3 = z$	9
(f) $3y = z$	36
(h) [no attempt]	10

Table 1 : Data from MacGregor & Stacey

Given that the problem is phrased in the same word order as a possible algebraic solution, our viewpoint would interpret that response (a) is a natural direct translation but (b) and (d) correspond either to flexible proceptual equivalents or to process orientation. Given the total lack of responses in (c), which represents the aesthetic order of linear equations, we infer that (b) and (d) are more likely process orientation rather than random proceptual flexibility. We would conclude that items (e) and (f) are more likely to involve primitive conjoining rather than mistaken multiplication. The number of responses in the last four (43%) show a horrendously large number of children who seem be seriously handicapped in learning algebra.

We designed an investigation in which three addition and three multiplication questions increased in difficulty, with the first two in each case having syntax coinciding with algebraic assignment order and the third was more complex so that the syntax did not easily suggest a specific order. The problems were given to two groups whom we hoped would be less likely to produce conjoining and non-responses. One was a group of 75 Year 9 students (age 13 to 14) in a highly selected school in Britain representing roughly the top 35% of the total population. The other was a group of 128 second year university students training to be teachers of children aged 4 to 12, representing the top 20% of the population overall. (These teachers are not mathematics specialists and the vast majority will not have studied mathematics for over three years.) The responses of three questions similar to those of McGregor and Stacey are given in Table 2, the codes **s**, **a**, **p**, **r**, ✔, ✘, standing for syntactic, assignment-oriented, process-oriented, reversal, correct and incorrect, respectively.

Item	**Response**		**School**	**University**
(1) *y* is equal to *x* plus four.	*y* =*x* + 4 [*y*=(*x*+4)]	**as✓**	65	107
	y = 4+*x*	**a✓**	0	0
	x + 4 = *y*	**p✓**	9	15
	4 + *x* =*y*	**p✓**	0	2
	y =*x* + 4 or *x* + 4 = *y*	**a✓**	1	3
	other	✘	0	0
	no response	✘	0	1
(2) *w* is equal to the sum of 3 and *n*	*w* = 3 + *n* [*w* =(3 + *n*)]	**as✓**	52	91
	w = *n* + 3	**a✓**	4	2
	3 + *n* = *w* [(3 + *n*) = *w*]	**p✓**	10	30
	n + 3 = *w*	**p✓**	2	2
	w = 3+*n* or 3+*n* = *w*	**a✓**	0	1
	other [e.g. 3*n*=*w*, *w*=3]	✘	7	2
	no response	✘	0	0
(3) A school has *v* girls and *t* boys. There are ten more girls than boys. Write an equation relating *v* and *t*.	*v* = *t*+10 [*v*=10+*t*, t=10–*v*]	**a✓**	41	33
	t+10 = *v* [10+*t*=*v*, *v*–10 = *t*]	**p✓**	24	44
	v+10 = *t* [10+*t*=*v*]	**pr✘**	7	24
	t = *v*+10 [*v*=*t*–10]	**ar✘**	1	4
	other [e.g. 10 *t*=*v*, *t*=10 *v*]	✘	2	20
	no response	✘	0	3

Table 2 : Translations of verbal addition problems into algebra

Note that, in the first two questions the vast majority follow the given word order *y*=*x*+4 and *w*=3+*n*. *None* reverse the right hand side of (1) as *y*=4+*x*, but a small number reverse the word order in (2) to give *w*=*n*+3 (the aesthetic order). There are also *no reversals* of the roles of the variables (the syntax does not encourage reversal). The majority of the remainder in both questions keep the word order (*x*+4 or 3+*n*) of the sum, but follow the procedural order. None of the students fail to answer question (1) correctly (it uses the familiar *x* and *y* in the correct order) but a small minority produce incorrect solutions to question (2) including conjoining. Notice too that the number reverting to the procedural solution increases as the problem becomes less familiar. The verbal problem (3) is more complex. There is no clear equation syntax and the solver must relate several different pieces of information in the mind at once – the number of girls (*v*), the number of boys (*t*) and deduce that the second sentence means *t* is 10+*v*. This will place stress on short-term working memory and there are more errors.

Tables 3 and 4 show the diminution in algebraic assignment and the increase in process-orientation as the complexity increases. Using a χ^2 test with continuity correction, the change in total process-oriented from (1) to (2) is statistically significant ($p<0.05$) amongst the university students and from (2) to (3) is highly significant ($p<0.01$) in both groups. When the first errors occur in (3), the process reversals exceed the assignment reversals. The difference is significant ($p<0.05$) calculated as a subset of the total school population, highly significant ($p<0.01$) of the corresponding university population. In addition, the proportion of those using the algebraic assignment making an error is smaller than the proportion of those using the process orientation. This difference is highly significant in the school students and significant in the university students.

Assignment oriented	**School (N=75)**			**University (N=128)**		
	All (as+a+ar)	**as✓+a✓**	**ar✘**	**All (as+a+ar)**	**as✓+a✓**	**ar✘**
(1)	65 (87%)	65	0	107 (84%)	65	0
(2)	56 (75%)	56	0	94 (73%)	94	0
(3)	42 (56%)	41	1	37 (29%)	33	4

Table 3 : assignment oriented notation and reversals in (1), (2), (3)

Process oriented	**School (N=75)**			**University (N=128)**		
	All (p+pr)	**p✓**	**pr✘**	**All (p+pr)**	**p✓**	**pr✘**
(1)	9 (12%)	9	0	17 (13%)	17	0
(2)	12 (16%)	12	0	32 (25%)	32	0
(3)	31 (41%)	24	7	68 (53%)	44	24

Table 4 :process oriented notation and reversals in (1), (2), (3)

Table 5 has three multiplicative questions increasing in order of difficulty. The simple multiplication problem (4) induces more errors than the addition problems (1) and (2) the majority are process-oriented in the form $5m=n$ which multiplies the first two items in the sentence (m and 5), but turns them into the accepted ordering $5m$. Question (5), though complex uses pop groups familiar to students and has the word order in the same order as a possible syntactic solution. Question (6) does not have a simple syntactic translation into algebra and causes even more errors.

Item	**Response**			**School**	**University**
(4) m is 5 times n.	$m = 5n$	[$m=n$x5]	**as✓**	49	89
	$5n = m$	[nx3$=m$]	**p✓**	19	26
	$5m = n$	[mx3$=n$]	**pr✘**	5	13
	$n = 5m$		**ar✘**	1	0
	other		**✘**	1	0
	no response		**✘**	0	0
(5) A record by Take That is h minutes long. A record by Kriss Kross is g minutes long. The Take That record is three times as long as the Kriss Kross. Write an equation relating h and g.	$h = 3g$	[$h=3$xg, $h=g$x3]	**as✓**	33	45
	$3g = h$	[3x$g=h$]	**p✓**	31	50
	$3h = g$	[3x$h=g$, hx3$=g$]	**pr✘**	9	25
	$g = 3h$	[$g=h$x3]	**ar✘**	1	5
	other		**✘**	1	3
	no response		**✘**	0	0
(6) A Band makes four times as many singles as albums. It makes q albums and z singles Write an equation relating q and z.	$z = 4q$	[$z=q$x4, $z=q$4]	**a✓**	20	13
	$4q = z$	[4xq=z, qx4=z]	**p✓**	23	59
	$4z = q$	[4xz=q, zx4$=q$]	**pr✘**	20	38
	$q = 4z$	[q=4xz]	**ar✘**	12	13
	other		**✘**	0	2
	no response		**✘**	0	3

Table 5 : Translations of verbal multiplication problems into algebra

Tables 6 and 7 show the decrease in algebraic assignment and the increase in process-orientation as the problems become more complex. Once again, as difficulty increases, the numbers in all of these categories increase. The change in total process-oriented from (3) to (4) is highly significant ($p<0.01$) in both groups and from (5) to (6) is highly significant amongst the university students. The changes in the process-reversal numbers

Assignment oriented	**School (N=75)**			**University (N=128)**		
	All (as+a+ar)	**as✓+a✓**	**ar✘**	**All (as+a+ar)**	**as✓+a✓**	**ar✘**
(4)	50 (7%)	49	1	89 (70%)	89	0
(5)	34 (45%)	33	1	50 (39%)	45	5
(6)	25 (33%)	20	5	26 (20%)	13	13

Table 6 : Assignment oriented notation and reversals in (1), (2), (3)

Process oriented	**School (N=75)**			**University (N=128)**		
	All (p+pr)	**p✓**	**pr✘**	**All (p+pr)**	**p✓**	**pr✘**
(4)	24 (32%)	19	5	39 (30%)	26	13
(5)	40 (53%)	31	9	75 (59%)	50	25
(6)	43 (57%)	23	20	97 (76%)	59	38

Table 7 : Process oriented notation and reversals in (1), (2), (3)

is significant from (4) to (5) in the university students and from (5) to (6) in both. The number of assignment reversals is always smaller than the number of process reversals, at a level which is significant in school in (4) and highly significant in all other cases. As proportions of those using the assignment order or process order, the errors on the latter are highly significant in (4), significant in school in (5), highly significant in university in (5), but not significant in either in (6). The last statistic simply states that, although the figures *look* different, this difference could have occurred by chance in more than one trial in twenty.

Discussion

As problem statements move from being syntactically equivalent to the algebraic assignment formulation to more complex statements, the student responses increase in their use of process-oriented statements with the operation on the left and the answer on the right. In cases where the syntax becomes too complicated to support a straight translation and the words are used in a way which does not encourage the use of letters as objects, errors which reverse the roles of letters still occur. Such errors may occur because of the cognitive complexity rather than a specific syntactic misconception.

A more detailed case analysis will be required to distinguish between those who think in a flexible proceptual manner and those who are more procedural. Flexible thinkers may respond in either the assignment form or a process-oriented equivalent, whilst the procedural child is likely to learn the algebraic assignment mode as a given procedure However, the regression to process-oriented representation with its higher level of failure is consistent with the fact that the method corresponding to earlier experiences in arithmetic is evoked in times of stress and proves to be more prone to failure. We see children responding to the translation process not only in purely syntactic word order, but also by attempting to make sense of the data using arithmetic and algebraic constructs related to their stage of development in algebraic sophistication.

References

Booth, L. R. (1984). *Algebra: Children's Strategies and Errors.* Windsor, Berks: NFER-Nelson.

Clement, J., Lochhead, J. & Monk, G.S. (1981). Translation Difficulties in Learning Mathematics, *American Mathematics Monthly,* 4, 286–290.

Dubinsky, E. (1993). Computers in teaching and learning discrete mathematics and abstract algebra. *Advanced Technologies in the Teaching of Mathematics and Science* (ed. D. L. Ferguson). Springer-Verlag, 525–563.

Foster, R. (in preparation): *It's that goat page!*, available from Math. Ed. Res. Centre, Warwick.

Gray, E. M. & Tall, D. O. (1993). Success and Failure in Mathematics: The Flexible Meaning of Symbols as Process and Concept, *Mathematics Teaching*, **142**, 6–10.

Gray, E.M. & Tall, D.O. (1994). Duality, Ambiguity and Flexibility: A Proceptual View of Simple Arithmetic, *The Journal for Research in Mathematics Education*, **26** 2, 115–141.

Gray, E. M. (1993). Count-on: The parting of the ways for simple arithmetic, *Proceedings of the Seventeenth International Conference of PME*, Tsukuba, Japan, 2, 204–211.

Herscovics, N. (1989). Cognitive Obstacles Encountered in the Learning of Algebra, In S. Wagner & C. Kieran (Eds.), *Research Issues in the Learning and Teaching of Algebra*, NCTM, Reston, Virginia, 60–86.

Kaput, J. J. (1987). Towards a Theory of Symbol Use in Mathematics, In C. Janvier (Ed.) *Problems of Representation in the Teaching and Learning of Mathematics*, Hillsdale NJ: Erlbaum, 159–195.

Kieran, C. (1981). Concepts Associated with the Equality Symbol, *Educational Studies in Mathematics*, **12**, 317–326.

Kuchemann D. E. (1981). Algebra. In K. M. Hart (Ed), *Children's Understanding of Mathematics : 11-16*, London: John Murray, 102–119.

Macgregor, M. & Stacey, K. (1993). Cognitive Models Underlying students' Formulation of simple Linear Equations, *Journal for Research in Mathematics Education*, 24 (3), 217–232.

Schoenfeld, A. H. (1985). *Mathematical Problem Solving*, New York: Academic Press.

Laborde, C. (1990). Language and Mathematics, in P.Nesher & J. Kilpatrick (Eds), *Mathematics and Cognition*, Cambridge: CUP, 53–69.

Mestre, J.P. (1988). The Role of Language Comprehension in Mathematics and Problem Solving. In R. Cocking & J. Mestre (Eds.), *Linguistic and Cultural Influences on Learning Mathematics*, Hillsdale, NJ: Erlbaum, 201–220.