The Psychology of Symbols & Symbol Manipulators

David Tall
Mathematics Education Research Centre
University of Warwick, COVENTRY CV4 7AL, UK

Symbol manipulators are reputed to "release the student from the drudgery of routine manipulation so that they can focus attention on the concepts and solving problems". Mathematicians have a "symbol sense" that enables them to handle symbols in a flexible manner, but students do not necessarily work in anything like the same way. Formally mathematicians seek precision and unique definitions, but *cognitively* they seem to use symbols *ambiguously* to represent either processes to *do* mathematics or concepts to *think* about. This has intriguing consequences for student learners using computers, particularly those whose mental imagery of symbolism is different from that of their teachers.

Symbols as Process and Concept: the notion of Procept

Following a long line of researchers, including Piaget (1952), Dubinsky (1991), Sfard (1991) and others, Gray & Tall (1994) introduced the notion of *procept* as an amalgam of three things—process, symbol, concept—to allow discussion of the phenomenon in which a process such as counting becomes a concept such as number, or a symbol such as $2+3x$ represents a process of evaluation and a concept of expression in algebra. We hypothesised that mathematicians use symbolism in an *ambiguous* way, to represent either process or concept as appropriate, but less successful thinkers may see mathematics as inflexible procedures, seeking the security of following a tried and tested route to obtain an answer in a limited context. Procedures which *take place in time* may allow students to *do* mathematics but they are less suitable for thinking *about* mathematics.

At all ages we found differences between flexible thinkers using symbols dually as process or concept and those relying on symbolism to cue routine procedures. In algebra, those who saw the symbols as procedures to be carried out are less likely to grasp the meaning of the symbolism. Students conceiving of $3+2x$ *as a process* do not see it making sense unless x is known to compute the value, but if x is known, there seems no reason to complicate matters by using the symbol x. An equation such as

$$5x+1=11$$

might make sense as a problem where five times a number plus one is eleven, so five times the number is ten, and the number is two. But the equation

$$5x+1 = 3x+5$$

would be less likely to make sense because the equals sign no longer means "makes" and there are now *two* processes to carry out, one on each side. The flexible thinker has a meaningful way of manipulating equations to obtain a solution, but the procedural thinker is more likely to learn mechanical routines (Tall & Thomas, 1991).

Further evidence can be gleaned from the way students interpret word problems algebraically. Crowley *et al* (1994) found that the more complicated the word problem, the more likely the student would write a "process" equation such as $x+4=y$ (in which "x plus 4 *makes* y") rather than the more standard "assignment" equation $y=x+4$ (y *equals* $x+4$). Furthermore, those writing the "process" version were more likely to make errors, showing that this is linked with students who are less successful.

At a higher level, limit concepts, such as $\lim_{x\to\infty}\frac{2x+5}{3x+4}$ or $\sum_{n=1}^{\infty} 1/n^2$ involve symbols which represent both a process of "getting close" and also the *value* of the limit. The research

As published in *Proceedings of the Seventh Annual International Conference on Technology in College Mathematics Teaching*, Addison-Wesley, (1995), pp. 453–457.

literature is full of examples of students who see the process "getting closer and closer" without actually "reaching" the limit, or perhaps "only reaching the limit at infinity" (summarised in Cornu, 1991). There seem to be "unavoidable misconceptions" (Davis & Vinner, 1986) and students cope by using localised procedures for individual problems as they occur (Williams, 1991).

Computer tools for manipulating symbols

In the previous discussion three essentially different kind of procepts arise:

- *Operational procepts*, such as those in arithmetic, or differentiation of symbolic formulae, with explicit algorithms to compute a result,

- *Template procepts*, such as expressions in algebra, containing variables where the procept may be evaluated by giving values to the variables, but may also be manipulated symbolically as mathematical concepts,

- *Structural procepts*, such as limits, representing a process to give a result but without a direct procedure to find it; instead a structure of the relationships may offer various other approaches.

Of these, *operational procepts* are the easiest to program. The four rules of arithmetic were came first, now readily available everywhere on hand-calculators. *Template procepts* are more complex and the simplification involves specifying a list of templates indicating which expressions (such as $x*(-y)$) may be replaced by others (such as $-(x*y)$) and performing an exhaustive search to simplify a given formula. Some cases, such as the solution strategy for certain simple equations may be programmed as an algorithm so that a symbol such as **solve(ax+b=cx+d,x)** becomes an operational procept instructing the computer to print the solution of the equation $ax+b=cx+d$ in terms of x.

Structural procepts need careful handling in symbol manipulators. Sometimes this involves an intellectual form of cheating. Whereas we might compute a limit such as

$$\lim_{x \to 0} \frac{\sin x}{x}$$

from first principles, perhaps using a visual argument, a symbol manipulator may have more success by using L'Hôpital's rule to get

$$\lim_{x \to 0} \frac{\sin x}{x} = \lim_{x \to 0} \frac{D(\sin x)}{D(x)} = \frac{\cos 0}{1} = 1 .$$

Limits may involve the student in potentially infinite processes allowing something to get close, but not equal to, a specific value. Computer algebra software may attack the problem using a totally different symbolic algorithm with a finite number of steps.

Symbol manipulators therefore use a variety of sophisticated programming devices internally to carry out required processes to reveal the result to the user. They may carry out the internal process without revealing the true relationship between process and concept embodied in the cognitive notion of procept.

Mental tools for thinking about symbols

The brain is a complex structure, with a huge long-term memory store and a limited focus of attention. To use it to its fullest extent requires *compression* of information to fit in the focus of attention and *conceptual links* to pull related ideas into the focus as and when required. Compression of information into manipulable symbols gives great power:

> I should also mention one other property of a symbolic system – its compactibility – a property that permits condensations of the order $F = MA$ or $s = \frac{1}{2}gt^2$. In each case … the semantic squeeze is quite enormous. (Bruner, 1966, p. 12.)

By writing a string of symbols and going through what Davis (1984) calls a "visually moderated sequence"—considering the symbols, manipulating, considering the new symbols, manipulating, … , and so on, until a solution is found—it is possible to use the short-term focus and long-term connections to solve problems symbolically.

If the concepts are not compacted sufficiently to fit into the focus of attention, or if the connections are inadequate, then a primitive strategy is to practise and routinize sequences of actions. The problem is that such procedural methods are limited in two ways. First they are *inflexible* and only operate in well-defined contexts (including carefully designed examination questions in a form students expect to answer). Second, a procedure cannot be conceived as an entity, other than "if I have *this* problem, then I use *that* procedure." Once it is started, a procedure must be carried through in sequence. If other procedures are required, they must precede or follow in sequence, so the student has great difficulty solving problems requiring two or more stages (Rashidi & Tall, 1992).

The fact that so many teachers see the necessity of teaching procedurally is not a vindication of the system, but an indictment of it. Certain student difficulties manipulating symbols may be papered over by using symbol manipulators to do the work, but the question must be asked whether the symbols have any meaning, other than as a procedural interface to input data for the computer to process and produce an answer.

Failure of the mind and computer to represent "real numbers"

Students have perceptions of decimals and the number line which conflict significantly with the formal definition (Romero & Ascárate, 1994). Whilst they are familiar with integers, rationals and finite decimals, they regard other special numbers such as π, e, √2 being different in kind, with infinite decimals "going on forever" and being somewhat "improper" (Monaghan, 1986). These images tend to be supported by the different types of number in a symbol manipulator without focussing on real numbers as elements of a complete ordered field.

Decimals cause further subtle problems. For instance, Wood (1992) found a sizeable minority of mathematics majors believed there is no smallest positive number (because half of it would be less), but there *is* a *first* positive number, namely, 0·000...01, corresponding to 1–0·999... Handling decimals, especially finite ones, seems to give a *discrete* sense to numbers, in strict order, increasing a digit at a time in the last place.

If you don't use it you may lose it!

Focusing on certain aspects and neglecting others may cause the neglected items to atrophy. For instance, students using *Derive* on hand-held computers to draw graphs of functions did not need to substitute numerical values for the independent variable to get a table of values to draw the graph. As a result, they had little practice of numerical substitution. This had unforeseen consequences. Some students who could calculate by substitution before the course were unable to do so afterwards. The students were asked:

"What can you say about *u* if *u*=*v*+3, and *v*=1?"

None of the seventeen students improved from pre-test to post-test and six successful on the pre-test failed on the post-test (Hunter, Monaghan & Roper, 1993).

New procedures for old

Symbol manipulators provide ways of solving problems using the software to perform the manipulations internally. But they do not remove the procedural aspects from the mathematics. Instead they introduce new procedures. For instance, using *Derive* replaces the procedure of symbolic differentiation by a sequence of keystrokes:

- select **Author** and type in the expression,
- select **Calculus**, then **Derivative**,
- specify the variable (e.g. **x**),
- **Simplify** the result.

When two groups of students in UK schools were asked:

> Please explain the meaning of $\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$,
>
> … all the students [following a standard course] gave satisfactory theoretical explanations of the expression but none gave any examples. However, none of the *Derive* group gave theoretical explanations and only two students [out of seven] mentioned the words 'gradient' or 'differentiate'. Four of the *Derive* group gave examples. They replaced *f*(*x*) with a polynomial and performed or described the sequence of key strokes to calculate the limit. (Monaghan, Sun & Tall, 1994.)

Failure of some computer approaches to the limit concept

Given students' (and professors'!) idiosyncratic view of real numbers, it is no wonder that they also develop idiosyncratic views of the limit concept. Programming in a computer language or numeric system with limited accuracy does not seem to improve matters. Li & Tall (1993) found that the increasing time taken to compute more terms of a series emphasised the limit as an unfinished *process*, not a *concept*. They also found that programming had little effect on the reading of equations from left to right, so that the majority continued to believe that $0\cdot 1+0\cdot 01+0\cdot 001+\ldots=1/9$ is *false* but $1/9=0\cdot 1+0\cdot 01+0\cdot 001+\ldots$ is *true*. If $0\cdot 1+0\cdot 01+0\cdot 001+\ldots$ is seen as a process, rather than a value, The first equation represents a potentially infinite process which can never be completed but the second shows 1/9 being divided out to get successive terms.

Symbol Manipulation and Mathematical Proof

Symbol manipulators focus on getting results rather than taking the user through the process of manipulation to obtain the result so that students may fail to experience the meaningful interplay between process and concept. Even more seriously, they are concerned with manipulations that *work*, not those which *faii*, so they are unlikely to provide appropriate images for non-examples of concepts such as non-differentiability or discontinuity. They therefore may fail to motivate essential theorems in analysis. Carefully designed visual support may be provided by drawing suitable graphs, for instance functions taking different values on rationals and irrationals (e.g. Rosenthal, 1992) but graphs usually involve prototypical ideas intimating restricted mental imagery.

Reflections

The enthusiasm for new and exciting technology often focuses our mind more on the mathematics than on the thinking of the student. But now the first wave of technological change is stabilising, it is time to subject new approaches to more careful scrutiny to find out what is *really* happening under the surface. As we have seen in the examples in this presentation, it is likely that any attempt to use the computer in mathematics learning will have gains and losses. Indeed, under the surface, the student's images of mathematics are very different from the mathematical formalism they are intended to embody. It is therefore right to focus our attention on the students' thinking processes and address the wider issues of what is happening in their learning.

References

Bruner, J. S. (1966). *Towards a Theory of Instruction.* Cambridge: Harvard.

Cornu, B. (1991). Limits. In D. O. Tall (Ed.). *Advanced Mathematical Thinking* (pp. 253–166). Dordrecht: Kluwer.

Crowley, L., Thomas, M. O. J. & Tall, D. O. (1994). Algebra, symbols and the translation of meaning, *Proceedings of PME XVIII*, Lisbon, II, 225–232.

Davis, R. B. & Vinner, S. (1986). The notion of limit: some seemingly unavoidable misconception stages. *Journal of Mathematical Behaviour,* 5 (3), 281–303.

Davis, R. B. (1984). *Learning mathematics: the cognitive science approach to mathematics education*. Norwood, NJ: Ablex.

Dubinsky, E. (1991). Reflective Abstraction. In D. O. Tall (Ed.), *Advanced Mathematical Thinking* (pp. 95–123). Dordrecht: Kluwer.

Gray, E. M. & Tall, D. O. (1994). Duality, ambiguity and flexibility: a proceptual view of simple arithmetic, *Journal for Research in Mathematics Education*, 26, 115–141.

Hunter, M., Monaghan, J. D. & Roper, T. (1993). The effect of computer algebra use on students' algebraic thinking. In R. Sutherland (Ed.), *Working Papers for ESRC Algebra Seminar*, Institute of Education, London.

Li, L. & Tall, D. O. (1993). Constructing different concept images of sequences and limits by programming. *Proceedings of PME XVII*, Tsukuba, Japan, II, 41–48.

Monaghan, J. D. (1986). *Adolescent's Understanding of Limits and Infinity.* Unpublished Ph.D. thesis, Warwick University, UK.

Monaghan, J. D., Sun S. & Tall, D. O. (1994). Construction of the limit concept with a computer algebra system, *Proceedings of PME XVIII*, Lisbon, III, 279–286.

Piaget, J. (1952). *The child's conception of number,* (C. Gattegno & F. M. Hodgson, Trans.). London: Routledge & Kegan Paul.

Rashidi Razali, M. & Tall, D. O. (1992). Diagnosing students' difficulties in learning mathematics, *Int J. Math Ed, Sci. & Techn*., 24 (2), 209–202.

Romero i Chesa, C. & Azcárate Gimenez, C. (1994). An enquiry into the concept images of the continuum trying a research tool, *Proc. PME XVIII*, Lisbon, II, 185–192.

Rosenthal, J. W. (1992). Seeing is Believing in Analysis? *Proceedings of the Fifth Annual International Conference on Technology in Collegiate Mathematics*. Addison-Wesley, 437–452.

Sfard, A. (1991). On the dual nature of mathematical conceptions: reflections on processes and objects as different sides of the same coin, *Educational Studies in Mathematics,* **22** 1, 1–36.

Tall D. O. & Thomas M. O. J. (1991). Encouraging versatile thinking in algebra using the computer, *Educational Studies in Mathematics*, **22**, 125-147.

Williams, S. R. (1991). Models of limit held by college calculus students, *Journal for Research in Mathematics Education*, **22** 3, 237–251.

Wood, N. G. (1992). *Mathematical Analysis: a comparison of student development and historical development*. Unpublished PhD Thesis, Cambridge University, UK.