Mathematical Processes and Symbols in the Mind

David Tall
Mathematics Education Research Centre
University of Warwick
Coventry CV4 7AL
U.K.

Introduction

Mathematics, as taught to students, is continually the subject of scrutiny to see if it is appropriate for its task. In particular, the calculus, for so long conceived as the essential foundation of college mathematics, is being questioned as to its value in the wider realms of advanced education. Students seem to find much of it so difficult, and now, in the new era of information technology, symbolic manipulators are available which can do much of the algorithmic work of the calculus, so the question becomes : if a computer can do the work, why force students to do it? If the computer is available, why not forge a new partnership between student and computer in which each contributes their special abilities to produce a greater whole?

This article will consider the nature of human thinking processes to see how symbolism is utilised in mathematical thinking and to consider how technology is best integrated into the education process. In particular it will look at what kind of thinking a good mathematician performs which seems to make the mathematics so much easier than that faced by the average, or below average, student. We shall see that the child growing into the adult faces problems at every stage which relate to the divergence between the thinking of the successful mathematicians and those who eventually fail. Regrettably so many of the latter persist into college level mathematics with a way of thinking that makes their method of doing mathematics so much harder for them than the mathematics performed by their professors.

The Growth of Mathematical Ideas

The human mind is the product of five million years (and more) of evolution. Yet the growth in mathematical knowledge is exponential with more new ideas being developed each year than have ever occurred before. The foremost renaissance thinkers could hope to be poets, philosophers, musicians, mathematicians and many other things besides. Today knowledge grows at such a rate that the expert mathematician can no longer hope to encompass the whole of mathematics, gaining expertise instead in a relatively small part of the total.

On the other hand, there is no reason to suppose that there is anything dramatically different about the fundamental human apparatus of thinking than was present say two thousand years ago at the height of Greek mathematics, or ten thousand years ago with prehistoric man. Yet we expect our average student to cope with a knowledge base beyond that available in totality to any previous generation. What is it that enables this growth in knowledge to be encompassed in the minds of ordinary mortals of today's generation? First it is through the use of *language*, that enables the communication of thought, and through *written symbolism* that enables the essence of this thought to be passed on from generation to generation. But what is more important still is the manner in which the underlying concepts develop and the way in which the symbolism is used to assist the development of these concepts.

An analysis of the evolution of mathematical ideas shows that different parts of mathematics involve different kinds of thinking processes. Classical Greek geometry arises from observations of properties of specific kinds of objects which are idealised as mathematical models: points, lines, triangles, circles. These properties are *described* in a general manner which allows constructions to be carried out in a specified way, for instance, drawing the three angle bisectors of a triangle and observing that they are concurrent. Then arises the desire to show that this will always be so, resulting in the concept of geometric *proof*. The *descriptions* of geometric objects need to be refocussed as *definitions* that prescribe the mental objects from which deductions can be made. There is a desire to refine the theory to make the definitions minimal (it is not necessary to say that a square has equal sides and four right-angles – with equal sides, *one* right angle will do). But the symbolism used here: letters for points, two letters for a line, three for a triangle, and so on, all stand for a mental idealisation of objects which exist in reality. The detachment from reality is more a matter of philosophy than fact: demonstrated at the end of the nineteenth century by the realisation that there remained a dependence on geometric actuality because concepts such as "between" or "inside" had yet to be formally defined, but were an implicit part of the theory.

Number and algebra are different. These involve *processes* which are eventually *symbolised* in such a way that the symbols act dually for both the *process* and the resulting *concept*. This sequence of process becoming concept has become a major focus of mathematics education research in recent years (Beth & Piaget, 1966; Greeno, 1983; Sfard, 1991; Harel & Kaput, 1991; Dubinsky, 1991; Gray & Tall, 1991). It underlies the fundamental growth of modern areas of mathematics:

arithmetic, algebra, calculus and analysis. It will play a crucial role in the successful use of symbolic manipulators in education.

Symbols representing both process and concept

Processes are carried out and represented by symbols which subsequently take on a dual role, evoking either the process itself or the product of the process, depending on the context. Thus it is that:

- 5+3 represents both the process of addition and the concept of sum,
- 5x3 represents both the process of multiplication (through repeated addition 5+5+5) and the concept of product,
- The symbol 3/4 stands for both the process of division and the concept of fraction,
- The symbol +4 stands for both the process of “add four” or shift four units along the number line, and the concept of the positive number +4,
- 3+5x represents both the process “add 3 to the product of 5 and x” and the concept of the algebraic expression,
- The function notation $f(x)=x^3-27$ simultaneously tells both how to calculate the value of the function for a *particular* value of x and encapsulates the complete concept of the function for a *general* value of x,
- An “infinite” decimal representation $\pi=3.14159\ldots$ is both a process of approximating π by calculating ever more decimal places and the specific numerical limit of that process,
- Various limit notations, such as:

$$\lim_{x\to a} f(x),\ \lim_{n\to\infty} s_n,\ \lim_{n\to\infty} \sum_{k=1}^{n} a_k\ ,\ \lim_{\Delta x\to 0} \sum_{x=a}^{b} f(x)\ \Delta x\ \text{, etc,}$$

represent both the process of *tending to a limit* and the concept of the *value of the limit*.

What makes mathematical thinking powerful is the flexible way in which this conceptual structure is used. By using the symbolism to evoke a process, it can be used to compute a result, and by thinking of it as an object, it can be used as part of higher level manipulation. This results in a tremendous *compressibility* of mathematical conceptions. A compact symbolism can represent a complex concept which may also be mentally manipulated as a single entity. This proves to be a powerful tool for the mathematician, though it may cause a barrier for the learner who lacks the flexibility of meaning.

> One finally masters an activity so perfectly that the question of how and why students don't understand them is not asked anymore, cannot be asked anymore and is not even understood anymore as a meaningful and relevant question. Freudenthal (1983; p. 469)

So it is beholden to us as educators, if not as mathematicians, to analyse this process of compressibility to formulate ways in which it might be made available to a wider range of student ability.

The amalgam of process and concept as "procept"

The flexible use of a symbolism as either process or concept, so freely available to the professional mathematician, causes great difficulty for many students. It is well-recognised (e.g. Harel & Kaput, 1991; Dubinsky, 1991) that the composite of two functions f, g, can be conceived process-wise in the notation g(f(x)): first calculate f(x) and then calculate g of the result. But if the composite function $g \circ f$ is to be considered as a mathematical object in its own right, given in terms of the mathematical objects f, g, then a good deal of mental movement from concept to process and back again becomes important.

> Initially, functions are processes and so the subject must have performed an encapsulation in order to consider them as objects. It is important, for example in composition of functions, for the subject to alternate between thinking about the same mathematical entity as a process and as an object.
>
> (Dubinsky, 1991)

The question we should ask ourselves is "do mathematicians *consciously* think always that they are alternating between thinking of a function as a process and an object?" I think not. Having compressed the ideas (through using symbols), we simply use the symbol to denote whichever mental representation is appropriate, often without realising consciously what we are doing.

In the minds of successful mathematicians a symbol evokes either process or concept, whichever is appropriate, and this is done so subconsciously that we may be unaware that it is happening. To allow this idea to be a focus of attention, my colleague Eddie Gray and I formulated the term "procept" to mean:

> ... the amalgam of process and concept in which process and product is represented by the same symbolism (Gray & Tall, 1991)

This is intended to allow us to focus on the fact that good mathematicians think of a procept in a way which exhibits *duality* (as process or concept), *flexibility* (using whichever is appropriate at the time) and *ambiguity* (not always making it explicit which we are using).

The ambiguous use of symbolism is seemingly anathema in mathematical formalism, where definitions are made which quite clearly formulate a concept in one specific form ("a function is a *set* of ordered pairs such that ..."). Yet, having *defined* a function as a set of ordered pairs, as a *thing*, we then blithely go on to use it *as a process* or *as an object* whichever suits us at the time. The question to be asked therefore is: by making this flexibility explicit, can we help students develop these kinds of thinking processes? Before responding to this, let us consider the different kinds of thinking process that occur in practice.

Procedural and proceptual thinking

Our research with students of different ages, from kindergarten to college, shows a surprising similarity of difficulty at all levels. In traditional mathematics it is necessary first to acquire the ability to carry out a procedure, and then, after long practice, this is compressed mentally into a more compact mental object, often through the use of appropriate symbolism, to enable the mental object to become the focus of attention at a more abstract level. Students (from an early school age up) initially see the task as conquering the procedure. The more able soon encapsulate the procedure by use of appropriate symbolism and develop a flexibility with the notions that enable them to derive new ideas from old. A child may not *know* the value of 4x7, but might think of it as "four sevens" and know that "two sevens are fourteen" so "four sevens are fourteen plus fourteen, which is twenty eight". This method of *deriving new knowledge from old* is a natural consequence of proceptual thinking. I claim that the *proceptual thinker has a* ***built-in knowledge generator***. It is not necessary for such an individual to work hard to get results, these results are an automatic product of the knowledge structure. I conjecture that this flexible proceptual mode of thought is a major factor in the ability of the more able to do mathematics seemingly with little effort. Such a structure is organic. With a little fertilization it grows naturally almost of its own accord.

The less able, on the other hand, are more likely to focus on the currently required procedure as the main aim of the task. Success for them is being able to carry out the procedure and produce the required answer. Gray (1991) observed that children responding to simple arithmetic tasks seek the *security* of being able to carry out the procedure, rather than the *flexibility* of being able to derive facts from other known facts. Procedurally oriented children are often quite creative in developing their own methods for carrying out procedures which lead to short-term success. But this can also lead to long-term failure as the personal method may fail to cope with more complex tasks that require encapsulation of the procedure as an object for higher order manipulation.

These early steps in mathematics lead to patterns of thinking that can cause problems in college mathematics. It is my belief, for example, that the difficulties that average college student has with algebra occur because of previous rule-bound approaches to the subject. When students do not understand what something *is*, at least they can get temporary success by becoming secure with procedures to *do* things with it. In the early steps of algebra they meet an algebraic notation which generalises arithmetic. But whereas the arithmetic symbolism of operations such as addition are linked to a procedure to carry out the process and get the answer, the algebraic symbolism seems more obscure. The symbol $10–5x$ represents the process of taking 5 times x from 10, but it is a procedure which cannot be carried out until x is known and, if x *is* known, why use algebra anyway when arithmetic will suffice? Algebraic symbolism violates many individual's innate understanding of mathematical symbolism which in arithmetic tells them what to do and signals how to do it. The syntax is strange: why is $2+3x$ not computed from left to right as 3+2, which is 5, times x ? When students begin to feel uneasy, they often seek security in manipulating symbols to get the right answers. Each new topic is solved by learning a new and often seemingly arbitrary rule, "do multiplication before addition", "do operations in brackets first", "do the same thing to both sides", "cross-multiply", "put over a common denominator", "change sides, change signs", etc, etc, (Tall & Thomas, 1991).

We believe that these difficulties with algebra carry through to college students, and that the need for immediate procedural success, if not complemented by meaningful use of notation in the early stages, can so easily lead to meaningless symbol-pushing guided by these arbitrary rules. In Britain the fluency in algebraic manipulation at 16 years old is diminishing, although problem-solving abilities with numbers seem to be improving. The initial introduction of differentiation using the symbolic calculation of limits, even for a simple function like x^3, is severely compromised because it cannot be assumed that the whole population taking the subject can simplify the expression $((x+h)^3–x^3)/h$. We believe that this will lead to serious problems at college and university which may not be helped by the use of symbolic manipulators, unless this is part of a concerted effort to give proceptual flexibility to the meaning of the symbolism.

At higher levels the same proceptual difficulties recur again and again. Consider, for example, the product of two matrices. The procedural thinker will see the product as a calculation of each entry of the result through looking along a row of the first matrix and down a column of the second matrix, multiplying together corresponding pairs of entries, and adding together the results. The procedural product of two matrices involves a great deal of process. The proceptual thinker will see that this can be represented symbolically as the product AB of matrices A and B and, by thinking of the matrices as single objects and the product AB as an object, can begin to conceive of higher level structures such as (AB)C=A(BC), A(B+C)=AB+AC, or that, usually, AB≠BC, and so on. For the procedural thinker, these relationships occur not at the manipulable object level, but at the procedural level, involving far greater detail, far greater cognitive strain, and far greater difficulty for a less powerfully structured mind. No wonder the more able succeed almost trivially, whilst the less able are faced with catastrophic failure.

Once again, if students are procedural in their thinking, then they are faced with greater difficulties than if they develop proceptual flexibility. The same phenomenon occurs in other topics, for instance, in the understanding of limits, where students initially think of $\lim_{n\to\infty} s_n$ as a *process* of approaching a limiting value. They are faced with new problems here. To calculate the limit of, say

$$\frac{4n+3}{2n+1},$$

they may conceive this as “what happens to the calculation when n grows large?”. A common suggestion is that “the 3 and the 1 become small, so the answer is roughly $\frac{4n}{2n}$, which is 2”. They may develop a genuine intuition which helps them solve problems in a personal manner unrelated to the formal definition, but it may be an idiosyncratic method which fails in another context.

The proceptual structure of the limit definition is quite different from that of, say, arithmetic, where new derived facts may be obtained from old by using the same arithmetic operations. In the formal handling of limit, the student must cope with a difficult definition with several quantifiers. It is an awkward calculation, given an ε>0 to find an N such that

$$n>N \text{ implies } \left|\frac{4n+3}{2n+1} - 2\right| < \varepsilon.$$

Instead a new and initially unintuitive method is adopted. First, show that the definition applies to the convergence of simple sequences, for instance when $s_n = c$ (a constant), then $s_n \to c$, and when $s_n = 1/n$, then $s_n \to 0$. (The proof of these eminently self-evident results often raises an eyebrow of suspicion from students who fail to understand them.) Next a general theorem is proved, to say that if certain sequences tend to certain limits, then the sum, difference, product and quotients of the same tend to the obvious corresponding limits. (This theorem is "obvious", but its proof, in terms of unencapsulated ε–N processes, imposes enormous cognitive strain for very little apparent gain.) Then this is quoted to show that

$$\frac{4n+3}{2n+1} = \frac{4+3/n}{2+1/n} \rightarrow \frac{4+0}{2+0} = 2.$$

Thus it is that the new types of procedure cause great difficulties. The idea that $\lim_{n\to\infty} s_n$ is both a number, and a process, and that to calculate the number requires transition to process, thence manipulation using known facts and a general theorem to get the required result, is a type of proceptual thinking that once again shuts out the procedural thinker who cannot encapsulate the limit definition as a concept through the dual meaning of notation.

It is in fact, far worse, as any teacher at college level will know. The full-blown formalism of definitions of limits and axioms for the real numbers (including completeness) requires the learner to construct the properties of these defined objects by logical steps. This construction must be performed in a mind which already contains images of these properties, linked not to the definition, but to the students previous experiences. The subtlety of performing such constructions when many of the results, as exemplified by the student's mental images, seem already to be known, causes great confusion to the majority. It is actually made worse by teaching which acknowledges this difficulty and tries to be "more informal" with the mathematics, for the division between formal necessity and informal knowledge then become even more blurred.

In the case of the limit concept, students have an intuitive dynamic imagery which is in some ways in conflict with the formal definition (Tall & Vinner, 1981; Cornu 1983). (For instance, they may believe that the sequence "gets closer" to the limit but "never reaches it".) Compounding this difficulty is the immense problem of manipulating several quantifiers in the formal definition. The procedural thinker who attempts to handle the definition of a sequence in a procedural manner is faced with so much detail that, once again, failure, if not inevitable, is highly likely.

Thus there is a qualitative difference between different kinds of thinking processes. In school, the proceptual thinker develops a natural knowledge structure that, of its nature, generates new knowledge with little effort, the procedural thinker seeks instead the security of being able to *do* the processes of the mathematics, which often remains the sole focus of their effort. At college there are additional layers of sophistication which make the division between success and failure an even greater chasm.

Traditional teaching techniques usually focus on the procedural side, with the short-term aim of being able to *do* the mathematics. Once the procedures are sufficiently routinized to be able to carry them out almost subconsciously, they become the possible focus of reflective thinking and encapsulation as manipulable objects. It is possible therefore for a more able procedural thinker eventually to begin to compress procedural knowledge and to be able to reflect on it to move towards the required encapsulation. But the large cognitive structure required to carry out the procedure as a process in time mitigates against this success for many students. It may be a case of "not being able to see the wood for the trees", the cognitive complexity of the process completely overwhelming the conceptual simplicity of the (as yet unencapsulated) concept.

Using the computer to develop proceptual thinking

How then, given the divergence between more succesful proceptual thinking and procedural thinking that is likely to fail, can we begin to address the growing divide? Though one might formulate a policy of making explicit the very things that the more able do implicitly (using the symbolism flexibly as both process and concept, linking together different aspects of the concepts in flexible ways), there is an inherent difficulty. If the processes are not encapsulated, then coordinating processes occurs in time, it involves more low level detail, and it imposes greater cognitive strain on an already stressed individual. Failure seems inevitable.

However, the new technology gives a new and powerful facility. If the procedure can be automated as a computer algorithm, and virtually all of them can, then it may be possible for the computer to carry out the procedure, relieving the individual of the cognitive stress of coordinating the detail, and allowing the individual to concentrate on the relationships between the mathematical objects produced by the process. At one time the individual may concentrate on the procedure and what that entails, without thinking about higher level relationships involving the product of the procedure. At another time the individual may use a computer to carry out the procedure and concentrate on the relationship between the concepts. I term this the *principle of selective construction* (Tall, in press).

Whilst the traditional approach dictates that familiarity with the procedure must come before reflection on its product, using the computer there is sometimes a choice as to which may be done first. Sometimes the product of the procedure, if meaningful, may be explored before the procedure is practiced and interiorized by the individual, sometimes the procedure may be practiced before the product is studied in detail. This therefore gives new possibilities for curriculum sequencing.

For instance, in Tall & Thomas (1991) we attacked the problem of giving meaning to algebraic notation by a combination of programming in BASIC to calculate the values of algebraic expressions for given values of the variables, to see how the symbolism had a certain consistency and that different looking expressions (such as $2*(x+y)$ and $2*x+2*y$) always gave the same numerical result. Because the computer was carrying out the process of calculation, the student could concentrate on the products and see that $2*(x+y)$ and $2*x+2*y$ are equivalent expressions although as processes of computation they are different. We used a simple piece of software that accepts standard algebraic notation (with implicit multiplication and powers given by superscripts) to allow the students to see that $2(x+y)$ and $2x+2y$ are likewise equivalent (figure 1).

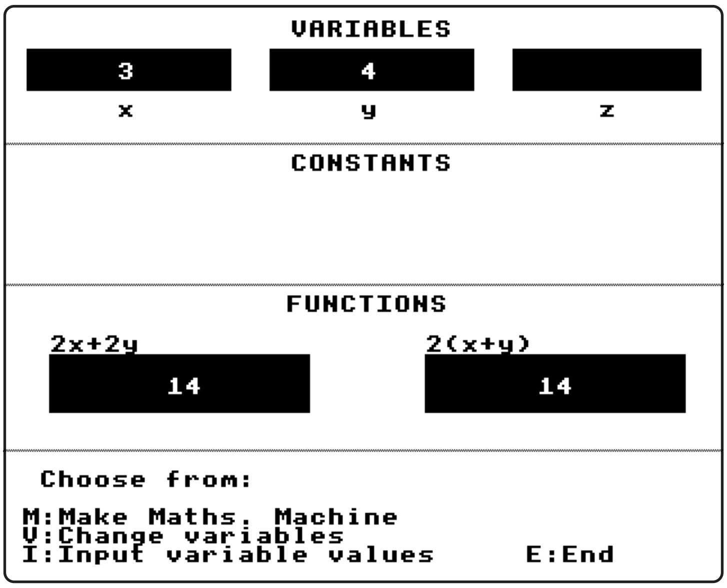

Figure 1 : the equivalent outputs of different processes

In Tall (1986) a similar experiment was done with the visual beginnings of the calculus in the English sixth form (Senior High School, ages 16-18), and showed a great improvement in the meaning of the derivative concept. By visualising the gradient of a graph as the gradient of a highly magnified (locally straight) small part of the graph, the student learned to look along a graph to *see* the changing gradient. The computer could draw an accurate representation of the numerical gradient $(f(x+h)-f(x))/h$ – carrying out a process that the student could not do with such precision – and the student could look at the graph of the gradient, see how it stabilized as h becomes reasonably small, and conjecture what the gradient curve might be approximating to. At college level, Heid (1984) has shown that a combination of conceptual learning at one stage using a symbol manipulator to carry out the routines of differentiation and later practice at the procedures of differentiation produced far more flexible and versatile learning.

The new calculus curriculum in the UK designed for 16 to 19 year olds by the *School Mathematics Project* uses software to visualise gradients and to guess the formulae of gradients of graphs before discussing the symbolic procedure of calculating limits. It also uses software to enable the learner to physically construct an approximate solution of a first order differential equation as the reverse process of knowing the gradient of a graph and using software to build up a curve which has the required gradient. This is done *before* considering considering any numerical or symbolic method of solution.

The differentials dx, dy are visualized as coordinates of the tangent vector to the curve y=f(x) in the x-y plane and in three dimensions the tangent to the curve $y=y(t)$, $x=x(t)$ has components (dt, dx, dy) (figure 2). This allows a differential equation in several variables to be given a physical meaning (as specifying the direction of the tangent) and allows formula such as

$$\frac{dy}{dx} = \frac{dy}{dt} \Big/ \frac{dx}{dt}$$

to be given a meaning as ratios of the components of the tangent vector. Figure 2 shows the tangent vector (calculated as a numerical approximation) drawn by the *Parametric Analyser* (Tall 1991).

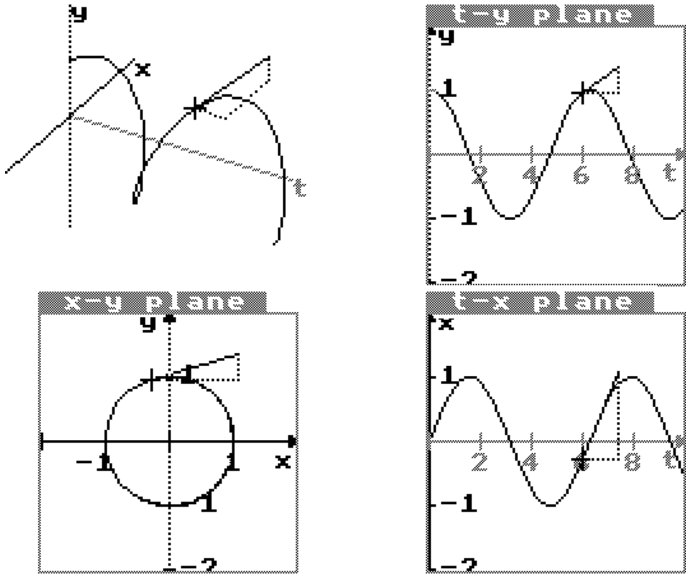

Figure 2 : The numerical tangent to the curve *x*=sin*t*, *y*=cos*t* in 3D
and the projections onto the coordinate planes

In these experiments we see visualization playing an increasing role in conceptual learning. In addition to the symbols representing both process and concept, they are also linked to other representations, such as graphical representations which expand the flexibility of conceptualization. The good mathematician evokes whichever representation is appropriate for a particular purpose, and uses that representation as long as it proves successful, switching to another representation when it proves more useful. This flexible form of thinking is often termed "versatile thinking" (Tall & Thomas, 1989). Robert & Boschet (1984) show consistently that the most successful students in advanced mathematics are those with the flexibility to work in more than one representation (graphic, numeric or symbolic). Those who are limited to one representation (usually numeric or symbolic) are less likely to solve a wide range of problems.

For example, if a student consistently relies on symbolic manipulation without other representations, how will that student respond to a symbol manipulator which gives the response:

$$\int_{-1}^{+1} \frac{1}{x^2}\,\mathrm{d}x = \left[-\frac{1}{x}\right]_{-1}^{+1} = -2.$$

The Achilles Heel of a Symbolic Manipulator in Education

Symbol manipulators were originally designed to solve the problem of programming the computer to do many algorithms in advanced mathematics. Initially it was far more a problem of getting the computer to jump through the appropriate hoops than it was of thinking of the eventual educational use of the software. Now symbolic manipulators are increasingly flexible and stable software environments, they are becoming more appropriate for teaching and learning as well as for research.

Initially manipulators were based on the teletype interface in which a line of symbols is typed and a response is given by the terminal. At the present time most remain environments where lines of symbols are input and evaluated, though the output may now include graphs as well as symbols (which may be printed in standard mathematical notation). Thus a line of symbols, conceived either as an expression to be evaluated or a process to be carried out, is given to the computer for processing. Essentially the software accepts a procept which it processes internally, performing a construction for the user and giving a response for the user to interpret. It performs one aspect of selective construction: that of carrying out the internal procedure.

More flexible interfaces are being developed offering other forms of communication. For instance, *Mathematica* allows a graph of a surface (initially input symbolically) to be pulled around, using a mouse to select the viewing point rather than requiring the coordinates explicitly. But the basic mode of communication remains symbolic input, internal processing and symbolic or graphical output.

What must be apparent from the discussion in the previous sections is that the use of manipulators demands a proceptual understanding of the symbolism involved. Thus, in education, the question is whether procedural thinkers can benefit from the use of symbolic manipulators, and whether the manipulator can be used in a wider educational context to promote more flexible proceptual thinking. The Achilles heel of the symbolic manipulator in education is the need for the individual to construct a meaning for the symbolism as flexible process and product and the fact that symbolic manipulators process input internally in a manner which may not be transparent to the user. The mere surface manipulation of symbols is not enough.

Given the worsening ability of students with algebraic manipulation (certainly in Britain where there is now more emphasis on numerical problem-solving than on algebra) the need to give *meaning* to the symbolism becomes even more important. In the UK teachers are finding that beginning calculus students are less likely to be able to find the local maxima and minima of cubics because, although they can differentiate the expressions, they cannot factorize the resulting quadratic. The latter could, of course, be performed trivially by a symbolic manipulator, but if the symbolism has little meaning, what use will this be?

If the students *are* able to give some meaning to the symbolism, then manipulators can be of definite assistance. For instance, in a course at college calculus level for student teachers at Warwick University, the accent was placed on visualizing the concepts first. The students knew the meaning of differentials dx and dy as components of the tangent vector. They could see, in a three dimensional picture, the meaning of partial derivatives through taking cross-sections of a surface and looking at the gradient of each. They could do simple differentiation and integration in a meaningful way. But when it came to finding the maxima and minima on a surface z=f(x,y), the sheer drudgery of working out second derivatives in certain examples defeated them. Here the use of a symbol manipulator (*Derive* and *Mathematica* were both used) proved to be greatly illuminating. The initial calculations are shown in figure 3. It proved a simple matter to get the software to calculate the second partial derivatives and check the required conditions.

The use of notebooks in *Mathematica*, which give electronic text whose symbols may be selected, evaluated, modified and investigated, promises greater flexibility for the active learner, although this still must be done within the syntax and facilities allowed by the software, with the internal procedures hidden from sight. But it should be remembered that the relationships generated by such manipulators work only in certain ways, for instance, from symbolic input to graphical output. Other directions, for instance, using graphical concepts to produce related symbolic notions, still need to be done by the human mind.

Using the principle of selective construction, one may hypothesise that the manipulator must be used as part of a wider educational context in which active learning is encouraged to develop flexible use of symbols as procepts and to link these symbols to other representations. This needs to be performed within a wider educational framework that encourages the student to develop a flexible understanding of the symbolism.

```
z=(a x + b y +c)^2/(x^2+y^2+1)

                2
(c + a x + b y)
----------------
      2    2
  1 + x  + y

p=D[z,x]

                                          2
2 a (c + a x + b y)   2 x (c + a x + b y)
------------------- - --------------------
        2    2               2    2 2
    1 + x  + y          (1 + x  + y )

q=D[z,y]

                                          2
2 b (c + a x + b y)   2 y (c + a x + b y)
------------------- - --------------------
        2    2               2    2
    1 + x  + y          (1 + x  + y )

r=D[z,x,x]

      2                                   2                2
   2 a        8 a x (c + a x + b y)   8 x  (c + a x + b y)
----------- - --------------------- + --------------------- -
     2    2            2    2 2                2    2 3
1 + x  + y       (1 + x  + y )           (1 + x  + y )

                    2
   2 (c + a x + b y)
   ------------------
          2    2 2
     (1 + x  + y )
```

Figure 3 : The initial calculations of a max/min problem with *Mathematica*

The need for versatile learning

Symbols alone cannot provide a total environment for mathematical thinking. They must represent something, and are more powerful if they do so in a flexible proceptual way. The power is further enhanced if there are alternative representations available which increase the flexibility of thinking.

Tall (1990) analyses the content of the calculus syllabus from a cognitive viewpoint, concentrating on the processes of differential and integral calculus. To my surprise, I found that the basic cognitive concepts were not differentiation and integration. Instead, I found that I needed to start with the more fundamental notion of *change* and see differentiation and integration as the symbolic parts of *rate of change* and *cumulative growth*.

The notion of change is represented by the *function* concept, which may now be seen as a flexible procept. It can be carried out as a process of assignment, it can be reversed as an inverse function, or as a solving of equations. It has several different representations, of which I concentrated on three: the symbolic, the numeric and the graphic. The symbolic and the numeric are both proceptual. They are procedures for calculating values of the function, which can also be conceived as objects (as expressions or as named computer procedures). The graphic representation (a function as a graph) occurs in a symbolic manipulator only as the output of a numerical procedure (which might be specified symbolically). As we saw earlier in our analysis of the difference between geometry and algebra, the visual concept tends to be seen as an object – a curve in space – rather than a process (take the value of x on the x axis, move up to the curve $y=\mathrm{f}(x)$ and across to the y-axis to find the corresponding function output). This is known to be a weakness of the graphical representation. But the graph also gives a large amount of qualitative information that enables the user to conceptualize global concepts that are often hard to imagine purely from the symbols or numbers. It therefore occupies a worthy place alongside symbolism and numeric procedures as representations of the fundamental notion of *change*.

Differentiation occurs as the symbolic part of *rate of change*, and integration as the symbolic part of *cumulative growth*. Each of these notions occurs as a process which can be done, and undone. The doing in each case is, of course, a procept, and the undoing is the reversal of the process part, which has a complementary proceptual structure. An interesting facet of this conceptual analysis is that the undoing of differentiation is *not* integration. It is the solving of a *differential equation*. In visual terms the qualitative idea of the derivative is the gradient of a graph, which may be seen by looking at the graph under high magnification so that is looks “locally straight”. It is possible at a primitive level to *see* the gradient of (the graph of) a function simply by casting ones eye along and estimating the changing slope. Once this is established and one can see a number of standard formulae (such as the derivative of x^2, x^3, x^n, $\sin x$, $\cos x$, $\ln x$, e^x, etc), the qualitative picture becomes an encumberance and one develops more powerful ways of calculating the gradient through rules of symbolic manipulation. Thus it is that the process of enaction of the gradient to be later replaced by the use of symbolic manipulation resembles earlier encounters with procepts. First it is necessary to give the concept a meaningful representation. Then aspects of the representation itself start to take on a life of their own (in this case the symbolic process of calculating a derivative) and the learner need only build on the meaningfully encapsulated procept (symbolically calculating the derivative), rather than stepping all the way back to first principles. The same is true of young children counting. At first this needs physical objects to operate on, but when the symbols have meaning, it is only necessary to depend on the meaningfulness of the symbols.

Solving a differential equation is the reverse of the process of finding the gradient: in primitive terms it is a matter of knowing the gradient everywhere and trying to build up the graph. Many pieces of software are available to draw direction diagrams and to draw numerical solutions automatically. The *Solution Sketcher* (Tall, 1991) allows the user to experience the physical act of building up a solution with the computer using the (symbolic) first order differential equation to calculate and draw a small line segment of the appropriate gradient through a selected point in the plane, (figure 4). Although this picture seems fairly innocuous, it is a potent enactive environment which enables the user to point anywhere in the plane and deposit a small line segment whose gradient is given by the differential equation. By putting such segments end to end and leaving a trace on the screen, the student can build up a solution curve. I see this enactive process of building up a solution curve as a fundamental physical action which gives a primitive meaning to the differential equation. Yet it forms the cognitive foundation of more formal concepts, such as the uniqueness of a solution through a given point (provided that the gradient is properly defined). Once the student has internalised the meaning of a solution of a differential equation in this way, it is soon apparent that the solution is

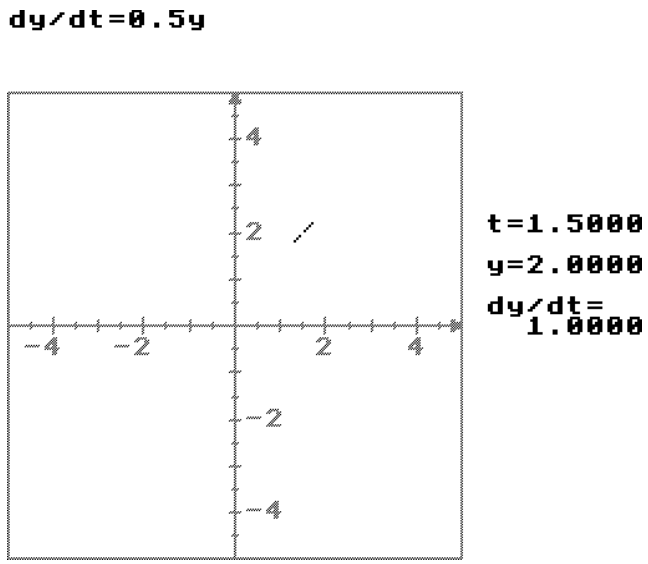

figure 4 : The Solution Sketcher
ready to build up the solution of a differential equation

found more easily either by using various numerical procedures, or through attempting to reverse the methods of symbolic manipulation.

Integration, as the area under a graph, or Riemann integration, may be classified as a cognitively different structure: *cumulative growth*. It has its own symbolic, numeric and graphic interpretations. Symbolic integration as a theory proves to be of great interest because of the fundamental theorem of calculus, which shows that it can be carried out by anti-differentiation. In practice the graphic and numeric interpretations of undoing cumulative growth are little studied because of the overwhelming power of the fundamental theorem.

This analysis of the structure of the calculus is given in figure 5. In this picture, the traditional part of calculus, the symbolic part, is but part of a much larger picture. This part is most valuable for its ability to carry out the procedures so successfully, where a picture might only give a qualitative idea and a numerical procedure only an approximation. But the symbolic undoing of differentiation, differential equations, involves many problems for which symbolic methods do not give a solution (in terms of elementary functions). Thus the wider picture, in terms of the practicality of solving problems and the visualisation the qualitative concepts, takes on an important overall role.

It should be very strongly emphasized that the existence of this structure does not mean that all the parts of the diagram need to be given equal weight, but that they should be used *for their appropriate purpose*. For instance, pictures should be used for conceptual insight, whilst numerical calculations or symbolic manipulation are used for productive calculation. The good mathematician selects whichever representation is appropriate for a given stage of a given problem, moving flexibly between representations where this becomes expedient. It is the versatility to move between representations and choose the most appropriate that gives the good mathematician great power. Such an approach does not overburden short-term memory by working on several different representations simultaneously. The desire to coordinate several different representations and to see processes carried out simultaneously in all of them can easily overstretch the working mental capacity. If it strains the good mathematician, it is even more likely to overburden the average student. What is more important is to allow the student to perform more in the mode of a good mathematician by allowing them to selectively construct part of the conceptual structure that is the current focus of attention whilst the computer carries out other parts of the constructive process.

Representations:		
Graphic	**Numeric**	**Symbolic**
Qualitative Visualizing Conceptualizing	**Quantitative** Estimating Approximating	**Manipulative** Formalizing Limiting

Concepts:

Change

doing:	graphs	numeric values	algebraic symbolism
undoing:	graphical solutions of equations – intersection of graphs	numerical solutions of equations – sequences of numerical approximations	inverse functions (solving equations) symbolic solutions

Rate of change

doing:	local straightness	numerical gradient of graph	derivative
undoing:	build graph knowing its gradient	numerical solutions of differential equations	solutions of differential equations –antiderivative

Cumulative Growth:

doing:	area under under graph	numerical area	integral
undoing:	know area –find curve	know area – find numerical function	FUNDAMENTAL THEOREM

Figure 5: The conceptual structure of the calculus

Summary

Symbolism is used flexibly by the good mathematician. Symbols allows mathematical thinking to be compressible, so that the same symbol can represent a process, or even a wide complex of related ideas, yet be conceived also as a single manipulable mental object. This flexibility is stock-in-trade for the mathematician. But it is not for the average student, who seeks a shorter-term goal: to be able to *do* mathematics by carrying out the necessary processes. It is this relationship between procedures to *do* mathematics and the encapsulation of these concepts as single objects represented by manipulable symbols that is at the heart of mathematical success and it is its absence which is a root cause of failure.

We therefore see that the use of symbols in a wider sense, dually representing processes or concepts, linked with other representations including visualisations, gives a flexible view of mathematics that makes the subject easier for the more able. The less successful tend to cling more to a single representation, often a procedurally driven symbolic approach, which is inherently less flexible and imposes greater cognitive strain on the user. The short term gain of showing a student the procedure to be able to *do* a piece of mathematics may, for these students, lead to a cul-de-sac in which security in the procedure prevents the flexible use of symbolism as both process (to obtain a result) and object (to be able to manipulate as part of higher level thinking). Now that computer environments are available to carry out algorithmic processes in a predictable manner, it may be possible to encourage a wider range of students to gain flexible insights into the higher level concepts, integrating them in a more proceptual manner, linking to other representations.

Symbolic manipulators, taking a proceptual input and internally carrying out procedures which are usually invisible to the user, may be used to complement the skills of the student, but this requires some insight into the meaning of the symbolism. Therefore the manipulators are better used as part of a richer environment which helps the students develop supportive linkages between concepts. They can provide an environment for manipulation of such symbolism, by carrying out the process and enabling the user to focus on the concept. This principle of selective construction offers a method of reducing cognitive strain and increasing the student's chances of developing more flexible thinking processes.

References

Beth, E.W. & Piaget, J., 1966: *Mathematical Epistemology and Psychology* (W. Mays, trans.), Dordrecht: Reidel.

Cornu, B. ,1981: 'Apprentissage de la notion de limite: modèles spontanés et modèles propres', *Actes du Cinquième Colloque du Groupe Internationale PME*, Grenoble, 322–326.

Dubinsky, E., 1991: 'Reflective Abstraction', in Tall D. O. (ed.) *Advanced Mathematical Thinking*, Dordrecht: Reidel.

Freudenthal, H. ,1983: *The Didactical Phenomenology of Mathematics Structures*, Reidel, Dordrecht.

Gray, E. M. , 1991: 'An Analysis of Diverging Approaches to Simple Arithmetic : Preference and its Consequences', *Educational Studies in Mathematics*, (in press).

Gray, E.M. & Tall D.O. ,1991: 'Duality, Ambiguity and Flexibility in Successful Mathematical Thinking', *Proceedings of PME XIII, Assisi*

Greeno, J., 1983: 'Conceptual Entities', in D. Genter & A. L. Stevens (eds.), *Mental Models*, 227-252.

Harel, G. & Kaput, J. J., 1991: 'The role of conceptual entities and their symbols in building advanced mathematical concepts', in Tall D. O. (ed.), *Advanced Mathematical Thinking*, Dordrecht: Reidel.

Heid, K., 1984: *Resequencing skills and concepts in applied calculus through the use of the computer as a tool*, Ph.D. thesis, Pennsylvania State University.

Robert, A. & Boschet, F. , 1984: 'L'acquistion des débuts de l'analyse sur ***R*** dans un section ordinaire de DEUG première année', *Cahier de didactique des mathématiques 7*, IREM, Paris VII.

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. , 1986: *Building and Testing a Cognitive Approach to the Calculus using Computer Graphics,* Ph.D. Thesis, Mathematics Education Research Centre, University of Warwick, UK.

Tall, D. O.,1990: 'A Versatile Approach to Calculus and Numerical Methods', *Teaching Mathematics and its Applications*, **9** 3 124-131.

Tall, D. O., 1991: *Real Functions and Graphs*, software for BBC compatible computers, Cambridge University Press.

Tall, D. O. (in press) : 'Interrelationships between mind and computer: processes, images, symbols', *Advanced Technologies in the Teaching of Mathematics and Science* (ed. David L. Ferguson), Springer-Verlag.

Tall, D. O. & Thomas M. O. J., 1989: 'Versatile Learning and the Computer', *Focus on Learning Problems in Mathematics*, **11** 2, 117–125.

Tall, D. O. & Thomas M. O. J., 1991: 'Encouraging Versatile Thinking in Algebra using the Computer', *Educational Studies in Mathematics*, **22** 2, 125–147.

Tall, D. O. & Vinner S. , 1981: 'Concept image and concept definition in mathematics with particular reference to limits and continuity', *Educational Studies in Mathematics*, **12** 2, 151–169.