Inconsistencies in the Learning of Calculus and Analysis

David Tall

University of Warwick
U.K.

Introduction

In the mid-seventies, as a lecturer in mathematics rather than a researcher in mathematics education, I first became concerned with the cognitive difficulties of students learning the calculus and analysis. At that time I believed, in common with other professional mathematicians, that the best way to help students is to present the materials in a logical and coherent manner. By preparing my lectures in this way I obtained the approval of both colleagues and students. The acid test of peer evaluation and consumer satisfaction gave this approach an extremely high rating.

But exploratory investigations into students' conceptions revealed fundamental inadequacies. Students learn to respond to standard questions in a predictable manner, but if their understanding is probed in unusual ways, subtle difficulties arise. We all see concepts through the rose-tinted spectacles of our experience and this colours our beliefs. Students bring pre-mathematical experiences into the classroom which affect their understanding of the mathematics. And the mathematics itself, though formalized into a coherent deductive system is based on implicit and explicit agreements between mathematicians which are not always totally consistent and may cause further difficulties.

In this paper I shall look at some of the reasons behind the inconsistencies in the learning of calculus and analysis based on my own research over the last decade and a half. It is helpful to focus on three different areas: the mind, the mathematics, and the message:

Published in *Focus.* **12** 3&4, 49–63 (1990)

The *mind* of the student (and also of the teacher and professional mathematician) with its idiosyncratic experiences and beliefs, and personal ways of building and testing ideas.

The *mathematics* : as developed and shared between the minds of mathematicians and practioners over the centuries - a shared theory that has the advantage over other theories that there are strong deductive relationships and coherences between concepts, nevertheless a theory that has belief structures (both explicit and implicit) about the nature of those concepts and the allowable ways in which they may be related.

The *message*: of mathematics as conveyed to students by teachers, textbooks and other media. Within the message are two important subcategories which should be emphasized: the *use of language* and the *sequence of presentation of ideas*.

Inconsistencies can arise in all of these:

(1) The *minds* of students, teachers and mathematicians have experiences and belief structures that are not always consistent and may assemble mathematical ideas in idiosyncratic ways,

(2) The *mathematics* contains concepts such as limit and infinity, which carry complex meanings that may be interpreted in inconsistent ways,

(3) The *message* may be framed in a *language* that evokes inappropriate ideas and may be presented in a *sequence* that is inappropriate for cognitive development.

These different strands are woven together in a complex web. A growing body of research has linked together (1) and (2), through the theory of concept image and concept definition (Tall & Vinner 1981) and more recent research has concentrated on addressing the problems of language and curriculum sequencing (see, for example, Tall 1986a).

The mind: inconsistencies identified in research

In Tall (1977) the results of the an investigation into students’ beliefs were reported, based on written responses to questionnaires by a population of 36 mathematics students in their first week of study at university.

One item made the request:

> If you know the definition of the limit of a sequence, write it down: $s_n \to s$ as $n \to \infty$ means:

A later one asked:

> Is $0{\cdot}\dot{9}$ (nought point nine recurring) equal to one, or is it just less than one? Explain the reason behind your answer.

Only 10 out of 36 students claimed to know a precise definition and only seven were able to formulate a definition that was mathematically acceptable. Of these seven, only *one* responded that $0{\cdot}\dot{9} = 1$.

Thirteen of the thirty six held apparently conflicting views, asserting that $0{\cdot}\dot{9}$ was less than 1, whilst elsewhere stating that

$$\lim_{n\to\infty}\left(1+\frac{9}{10}+\frac{9}{10^2}+\ldots+\frac{9}{10^n}\right) = 2.$$

A week later the students were asked to write down various decimals as fractions, including

> $0{\cdot}25$
> $0{\cdot}05$
> $0{\cdot}3$
> $0{\cdot}333\ldots$
> $0{\cdot}\dot{9} = 0{\cdot}999\ldots$

Two thirds of the students (24) now said that $0{\cdot}\dot{9} = 1$ (or 1/1), including 13 who had previously affirmed the result was less. Their written answers also exhibited the conflict in terms of crossings out and added comments.

Subsequent research by one of my Ph.D. students, Monaghan (1986), working with 16/17 year-olds studying the calculus, showed that “recurring decimals are perceived as dynamic, not static, entities and are not *proper* numbers. Similar attitudes exist towards infinitesimals when they are seen to exist”. Comparing students taking a calculus course with students of a similar ability who were not, he concluded that “the first year of a calculus course has a negligible effect on students’ conceptions of limits, infinity and real numbers”.

During the late seventies and early eighties many examples of such conflicts were noted across a wide range of topics, including secants tending to tangents (Orton

1977), verbal and other difficulties with decimals, (Tall 1977), real numbers and limits (Schwarzenberger and Tall 1978, Tall 1980b, Cornu 1981), geometrical concepts (Vinner & Hershkowitz, 1980), the notion of function (Vinner 1983, Sierpińska 1985a, 1988), limits and continuity (Tall & Vinner 1981, Sierpińska 1987), convergence of sequences (Robert 1982), limits of functions (Ervynck, 1981) the tangent (Vinner 1983, Tall 1987), the intuition of infinity (Fischbein et al 1979, Tall 1980c, 1981, Sierpińska 1987), infinitesimals (Tall 1980a, 1981, Cornu 1983), the meaning of the differential (Artigue 1986), and so on.

The mathematics : different mathematical paradigms

As a mathematician I had always believed in the universal truth of the mathematics that I had learned and went on to teach students. It came as something of a brutal shock to realize that, whilst there was a considerable proportion of sound deduction and reasoning, there are also arbitrary belief structures that give mathematics a relative rather than an absolute truth. The Greek dichotomy between time and space being either made up of indivisibles or being potentially infinitely divisible still remains in our culture today. Standard mathematics banished infinitesimals with the arithmetization of analysis in the second half of the nineteenth century, but infinitesimal conceptions remain deeply embedded in the mathematical psyche so that a popular Encyclopaedia of Mathematics, (West et al., 1982) can describe calculus as:

> ... a branch of higher mathematics that deals with variable, or changing, quantities ... based on the concept of infinitesimals (exceedingly small quantities) and on the concept of limits (quantities that can be approached more and more closely but never reached).

The work of Robinson (1966) gave a logical basis for the notion of infinitesimal but, instead of bringing about a revolution in which infinitesimals were accepted as Robinson hoped, it was resisted by the mathematical community which largely continued to use standard methods. Yet, embedded in our teaching of epsilon-delta analysis are notions of “arbitrarily small” quantities which Cornu (1983) showed were causing students to have beliefs in infinitesimals.

The child is the father of the man. The student with infinitesimal notions passes through a training in standard analysis, but conceptions of the infinitely small are merely supressed, not eradicated, and lie there, dormant, ready to be evoked at a later time.

Concept image and concept definition

The distinction between the ideas related to a concept as evoked in the mind and the form of words used to (attempt to) define the concept is the focus of attention in Tall & Vinner (1981). The phenomena are here interpreted in terms of the theory of *concept image* and *concept definition*, defined as follows:

> We shall use the term *concept image* to describe the total cognitive structure that is associated with the concept, which includes all the mental pictures and associated properties and processes. ... As the concept image develops it need not be coherent at all times. ... We shall the portion of the concept image which is activated at a particular time the *evoked concept image*. At different times, seemingly conflicting images may be evoked. Only when conflicting aspects are evoked *simultaneously* need there be any actual sense of conflict or confusion.

On the other hand:

> The *concept definition* [is] a form of words used to specify that concept.
> (Tall & Vinner 1981, page 152)

The paper considers the curriculum studied earlier by pupils in an attempt to postulate reasons for the mismatch between students' evoked concept images and their knowledge of concept definitions of limits of sequences, limits of functions and continuity. For example, the students had a concept image of a continuous function, which could have come from a variety of sources, not least being the colloquial meaning of the term in phrases such as "it rained continuously all day" (meaning there was no break in the rainfall). This viewpoint is often reinforced by teacher's attempts to give a simple insight into the notion of continuity by speaking of the graph "being in one piece" or "drawn without taking the pencil off the paper", thereby confusing the mathematical notions of continuity and connectedness.

A questionnaire administered to 41 first year university mathematics students (Tall & Vinner 1981) included the a question to investigate the students' concept images (figure 1).

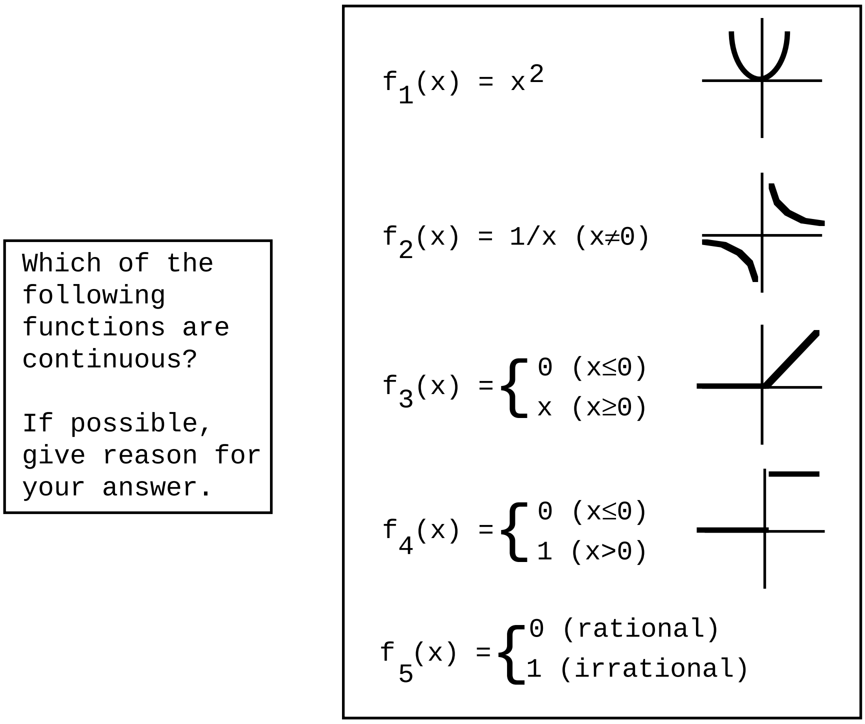

Figure 1 : the concept image of continuity

Mathematically f_1, f_2 and f_3 are continuous, whilst f_4 and f_5 are not. But the students' concept images suggest otherwise (Table 1 – "correct" responses in bold print).

N=41	f_1	f_2	f_3	f_4	f_5
continuous	**41**	**6**	**27**	1	8
discontinuous	0	35	12	**38**	**26**
no response	0	0	2	2	7

Table 1

Although all the responses to f_1 are "correct", the majority are "right answers for wrong reasons", such as the idea that f_1 is continuous "because it is given by only one formula". The function f_2 often causes dispute even amongst seasoned mathematicians. It is continuous according to the ε–δ definition *on the domain* $\{x \in \mathbb{R} \mid x \neq 0\}$. But the students' concept images suggest:

It is continuous:
because "the function is given by a single formula"

It is not continuous
because "the graph is not in one piece"

"the function is not defined at the origin"
"the function gets infinite at the origin".

Thus we see the concept image being evoked to respond to the question rather than the concept definition, leading to inconsistent responses.

It is possible for students to give apparently correct responses having evoked incorrect concept images, or to give glaringly incorrect responses for subtle reasons that have a large portion of truth in them. Tall (1986a) asked students to respond in writing whether they thought the statement in figure 2 is true or false.

```
As B→A the line
through AB tends
to the tangent AT.

True/False ? .....
```

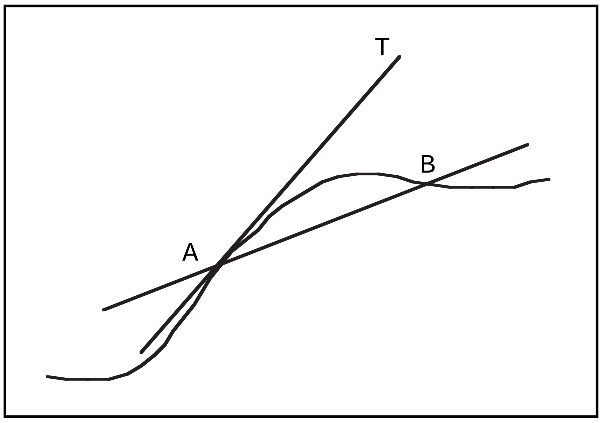

Figure 2 : tending to the tangent

Of a sample of nine 16 year old students interviewed in depth (as part of a larger project), four said the statement was "true" but linked the symbol B→A to vector notation and visualized B as moving to A *along the line BA*, so that the line (segment) BA "tends" to the tangent. Meanwhile, one student considered the statement "false" for the sensible reason that the lines were infinite so, way off at infinity, the line AB and tangent AT were still a long way apart no matter how close A and B became : an "incorrect" response for a very sensible reason. For this reason questionnaires alone, without follow-up interviews, may not reveal the full story.

The message: linguistic considerations

It is clear from what has already been said that language plays a large part in carrying both the meaning and evoking concept images that may be inappropriate and cause conflict. In a recent investigation I asked sixteen year old pupils to explain what a function is. The responses included the following:

- An equation with a variable factor – tells us what happens to a variable factor. e.g. f(x)=x+2.
- A process which can be performed on any number and is represented in algebraic form using x as a variable.
- An order which plots a curve or straight line on a graph.
- A series of calculations to determine a final answer.
- A term by which a sequence of numbers can be written and values calculated.
- A set of instructions that you can put numbers through.

Notice the many uses of technical mathematical terms in contexts where they have a colloquial rather than a technical meaning : variable, factor, order, series, term, sequence, set ...

It is quite clearly impossible to participate in human communication without such colloquiallisms. Thus the teacher, speaking to the student may use words with a technical meaning which the students interpret with a different colloquial meaning, or may use words colloquially in a manner which clouds the technical meaning. In my own teaching I have been made cruelly aware of this phenomenon. For example, I could not understand where my students obtained the idea that "a constant function is not really a function" – they considered y=4 "not to be a function of x because there was no x in the equation". I was quite certain that I had never said such a thing, yet on another occasion I found myself writing that the solution of a differential equation dy/dx=F(x,y) would have solutions differing by a constant "if the right-hand side is a function of x alone, and *not a function of y*". (The italics are here added for emphasis.) In a different mathematical context I had evoked the very concept I considered to be incorrect .

The message: curriculum sequencing

The sequencing of topics in the mathematics curriculum is built upon the implicit assumption that simple ideas must be introduced before more complicated ones – after all, this principle is so obvious, it is self-evident. But is it? The implications in the teaching of calculus might be that one should first experience calculus concepts with simple functions – polynomials and trigonometric functions – before leading on to more complicated functions such as those which are everywhere continuous and nowhere differentiable. This is certainly the way in which it is currently taught.

But this leads to all sorts of incorrect or inadequate beliefs which are common not only in students but also in many teachers: that a function must be given by a formula (and only one formula is allowed), that every function is differentiable, except possibly at a few isolated points, that the graph of a function looks fairly smooth with reasonably shaped maxima and minima, that graphs always have tangents, that a tangent touches the curve at one point only and does not cross the graph, etc etc.

It is my belief that we do students a disservice by organising the curriculum so that they are presented only with simple ideas first and given too great an exposure to an environment which contains regularities that do not hold in general. This just sows the seeds for later cognitive conflict. For example, doing geometry of curves only with circles, can give a dangerously limited idea of a tangent, or studying differentiation initially only with polynomials for so long may cause students to abstract general ”rules“ which are not true in a wider context: for instance that at a maximum the derivative is zero, rather than *if* the derivative exists, *then* it is zero...

Addressing the problem of cognitive conflict in mathematics education

We have seen a number of different sources of cognitive conflict: in the idiosyncratic meaning students and teachers give to concepts, in the subtle difficulties and inconsistencies in the mathematics, and in the way we attempt to transfer the message to students, through the ambiguities of language and the problem of developing appropriate learning sequences. We should therefore consider how we might address these problems to give teaching and learning strategies that face squarely up to the difficulties of cognitive conflict.

From the subtle and deep manner in which the use of language is intertwined in our thought, it becomes clear that we cannot improve matters simply by giving better concept definitions. That was an approach tried in the new mathematics of the sixties and, by and large, it failed. It failed because what might be an appropriate foundation for a logical mathematical development may not be an appropriate starting point for a cognitive development. Whilst we should be careful to present mathematics in a clear and consistent way we should not expect that the meaning we wish to convey will automatically transfer to the students.

On the contrary, what we must try to do is to provide students with learning experiences that will help them construct their own concept images, in a way which has sufficient richness to give better intuitions as to the likely truths in mathematics. Language alone cannot do this. But the computer affords a new resource which allows other forms of communication, via dynamic processes and visualizations which may be programmed to enable students to freely explore the concepts with the aim of gaining a greater insight.

There are two possible ways of coping with conflict (which are not mutually exclusive): one is to research the cognitive conflict to be prepared to face it when it occurs, a second is to give a richer conceptualization from the start to reduce the later conflict, or at least give the experience to set it in context. My own approach is to design computer software to provide a rich environment in which concepts could be demonstrated, explored and discussed. My thesis is that an environment allowing the user to explore both *examples* and *non-examples* of a mathematical concept or process can help the user abstract the general properties embodied in the examples and contrasted by the non-examples. An environment designed with this in mind is called a *generic organizer*, and (Tall 1986b) consists of a collection of generic organisers for visualizing calculus concepts. A new "graphic" approach to the calculus using these organisers is described in a series of six articles in Mathematics Teaching, starting with Tall (1985a). Powerful general ideas can be introduced at the outset (such as differentiable and non-differentiable functions at the beginning of the calculus) and specific examples with pertinent properties can be investigated (such as the tangent to a straight line or at an inflection point) to help students avoid narrow over-generalization. Six such generic organizers are:

- the magnify program: to magnify a tiny portion of a graph to investigate examples and non-examples of "local straightness" (i.e. examples of differentiable and non-differentiable functions).

- the gradient plotter: which has a secant through two close points on the graph move along the graph, simultaneously plotting the gradient of the secant as a point, dynamically building up the gradient function

- the area calculator: to show how the area under the graph of a function may be calculated as the area of a number of strips, or as a cumulative area function

- the solution sketcher: to enable the user to *construct* one or more solutions of a differential equation dy/dx=F(x,y) by drawing a short line segment of gradient F(x,y) at any point (x,y) and building up a solution by putting such segments end to end to give an approximate solution curve (figure 3).

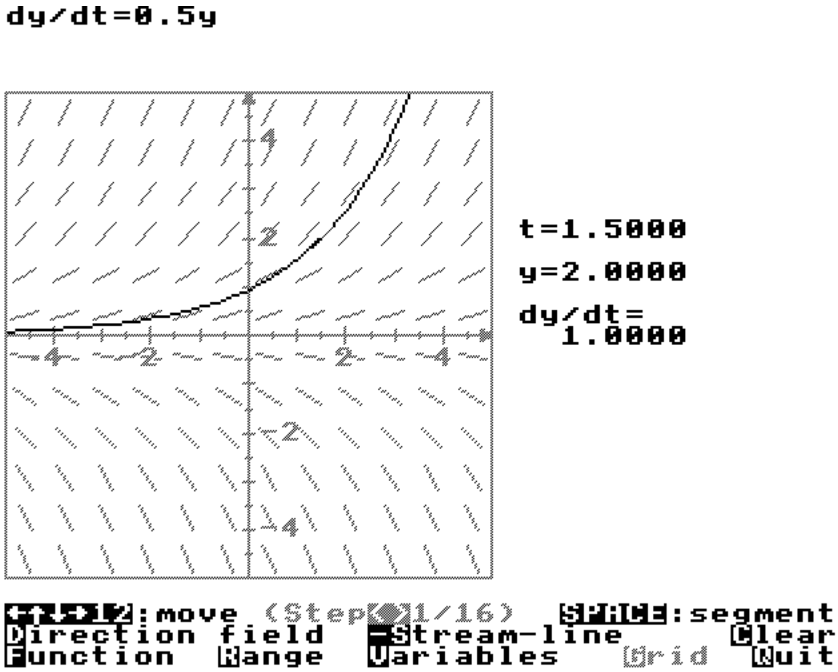

Figure 3 : The Solution Sketcher

- the three dimensional solution sketcher: to show simultaneous differential equations dx/dt=f(x,y,t), dy/dt=g(x,y,t) follow the same idea, with solution curves following the directions specified by the equations

- the parametric function analyser, enabling a three-dimensional view of parametric functions x=x(t), y=y(t) to be drawn, together with their projections on t-x, t-y and x-y space and to draw an approximate tangent vector and its components dt,dx,dy, to see that formulae such as

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

are just ratios of lengths of the components of the tangent vector (figure 4).

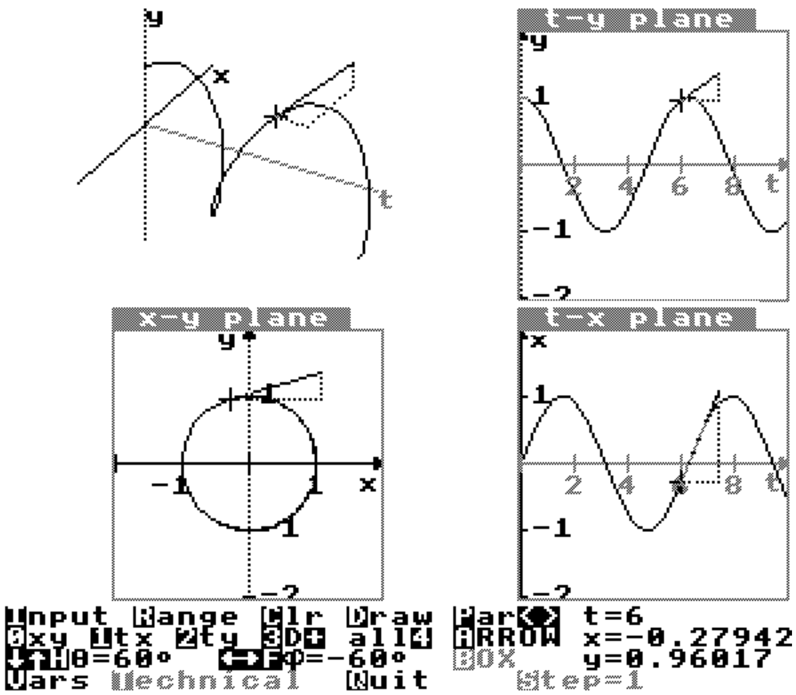

Figure 4 : The Parametric Function Analyser

All of these programs have been subsequently developed to allow extremely nasty functions (such as extremely wrinkled graphs to represent non-differentiable functions). Thus it is possible to investigate specific cases which are far more complicated than mere polynomials or smooth combinations of standard functions and give a much richer experience of *why* things can “go wrong” as well as surprising cases where they “go right” (Mills & Tall 1989). It is now possible to illustrate ideas that functions might be non-differentiable, or that the area function for a continuous, nowhere differentiable function *is* differentiable, and many more essential ideas that build experiences to give greater insight into analysis.

Computer software can provide a representation or model of the mathematical phenomena, but not always an exact translation. (For instance, pictures are drawn using finite pixels, so that straight lines do not normally look straight on a computer screen.) However, these obvious inconsistencies may be regarded as an *advantage*, not a hinderance. If the student can clearly *see* that the representation is not exact, it is possible to discuss the reasons why, and to begin to build up richer mental models.

The evidence from a number of studies so far (Tall 1986a, Blackett 1987, Thomas 1988) shows that the use of generic organizers in a context where the concepts are introduced by the teacher, discussed with the pupils and then explored by the pupils alone, suggests great advantages over traditional approaches.For instance Tall (1986a) built a “cognitive approach to the calculus” using the magnification and gradient drawing programs mentioned above. There was significant improvement in the experimental students’ ability to sketch gradients of given graphs, and their conceptualizations transferred to the more general case of the gradient and tangent of a graph at a point where the function was given by different formulae on either side.

Such an approach does not eliminate cognitive conflict. How could it? It is part of the human condition. But it does provide an open forum in which the conflict can be brought to the fore and discussed dispassionately through shared phenomena on the computer screen, instead of focussing on the personal hidden recesses of the pupil’s mind. For instance, the difficulties that Vinner (1983) observed with the student’s concept image of a tangent being a line which “touches a graph but does not cross it” was attacked using a graph plotter to draw lines through close points on a curve, producing a close approximation to a tangent. Zooming in close on the curve showed some graphs were “locally straight” and the tangent and graph were indistinguishable in the pixellated computer picture. Discussion of this point and what happens at points of inflection, or at a “corner” on a graph led to a much more flexible idea of a tangent and showed a significant change for the better (Tall 1987). New conflicts arose – for instance the students using the computer were, naturally, more likely to say that a tangent is a “line through two close points on a graph” because this is the way the concept was approached using the computer. But such concepts were more amenable to discussion and personal reconstruction by the pupils.

A new approach using computers in this way will need a radical reform of the curriculum. It applies at all levels of development. Much research, based on pre-computer environments, may be in error because it occurred in a context which may not pertain in future. For example, we teach young children about simple fractions in terms of halving and quartering because this is within their physical and mental capacity. But just because they may be *physically* incapable of dividing a cake into seven equal pieces does not mean that they are *mentally* incapable of visualizing it aided by appropriate software. A rich computer environment allowing children to carry out their mental ideas may give the opportunity to circumvent some of the trivializing introductions that may stunt future growth. We cannot make the complicated concepts more simple, but we can give far richer experiences that enable them to be seen in a wider, and more powerful, context. Recent research shows this to be a promising direction to follow (Tall & Thomas 1988).

References

Artigue M. 1986 : ‘The notion of differential for undergraduate students in the sciences’, *Proceedings of the Tenth International Conference of P.M.E.*, 229-234

Blackett N. 1987: *Computer Graphics and Children's Understanding of Linear and Locally Linear Graphs*, unpublished M.Sc. Thesis, The University of Warwick.

Cornu B. 1981 : ‘Apprentissage de la notion de limite : modèles spontanés et modèles propres’, *Actes due Cinquième Colloque du Groupe International P.M.E.*, Grenoble 322-326.

Cornu B. 1983 : *‘Apprentissage de la notion de limite: conceptions et obstacles'*, Thèse de doctorat, Grenoble.

Ervynck G. 1981: ‘Conceptual difficulties for first year university students in the acquisition of the notion of limit of a function’, *Actes due Cinquième Colloque du Groupe International P.M.E.*, Grenoble, 330-333.

Fischbein E., Tirosh D., Hess P. 1979: ‘The intuition of infinity’, *Educational Studies in Mathematics*, 1, 491-512.

Mills J.S.T. & Tall D.O. 1989: ‘From the Visual to the Logical’, *Bulletin of the I.M.A.* (to appear).

Monaghan J.D. 1986: *Adolescents’ understanding of limits and infinity*, unpublished Ph.D. thesis, Warwick University.

Orton A. 1977: ‘Chords, secants, tangents and elementary calculus’, *Mathematics Teaching*, 78, 48-49.

Robert A. 1982 : *‘L' acquisition de la notion de convergence des suites numériques dans l'Enseignement Supérier‘*, Thèse de Doctorat d'Etat, Paris VII.

Robinson A. 1966: *Non-standard Analysis*, North Holland, Amsterdam.

Sierpińska A. 1985a: "Obstacles epistémologiques relatifs à la notion de limite", *Reserches en Didactique des Mathématiques*, 6.1, 5-67.

Sierpińska A. 1985b: "La notion d'un obstacle epistémologique dans l'enseignement", *Actes de la 37e Rencontre CIEAEM*, Leiden.

Sierpińska A. 1987: "Humanities students and Epistemological Obstacles Related to Limits", *Educational Studies in Mathematics*, 18,4, 371-87.

Sierpińska A. 1988: "Epistemological remarks on functions", *P.M.E. XII, Hungary*, 568-575.

Schwarzenberger R.L.E. and Tall D.O. 1978: 'Conflicts in the learning of real numbers and limits', *Mathematics Teaching*, 82, 44-49.

Tall D.O. 1977: 'Conflicts and catastrophes in the learning of mathematics', *Mathematical Education for Teaching* 2,4 2-18.

Tall D.O. 1980a 'Intuitive infinitesimals in the calculus', *Abstracts of short communications, Fourth International Congress on Mathematical Education*, Berkeley, page C5.

Tall D.O. 1980b 'Mathematical intuition, with special reference to limiting processes', *Proceedings of the Fourth International Congress on Mathematical Education*, Berkeley, 170-176.

Tall D.O. 1980c 'The notion of infinite measuring number and its relevance in the intuition of infinity', *Educational Studies in Mathematics* 11 271-284.

Tall D.O. 1981 'Intuitions of infinity', *Mathematics in School*,10,3 30-33.

Tall D.O. 1985a 'Understanding the calculus', *Mathematics Teaching* 110 49-53.

Tall D.O.1985b 'Using computer graphics as generic organisers for the concept image of differentiation', *Proceedings of PME9*, Volume 1, Holland, 105-110.

Tall D.O. 1986a 'Building and Testing a Cognitive Approach to the Calculus using Interactive Computer Graphics', Ph.D. Thesis University of Warwick.

Tall D.O. 1986b *Graphic Calculus*, Glentop Press, London.

Tall D.O. 1987 'Constructing the concept image of a tangent', *Proceedings of the Eleventh International Conference of P.M.E.*, Montreal, III 69-75.

Tall D.O & Thomas M.O.J 1988: 'Versatile Learning and the Computer', *Focus*, (to appear).

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

Thomas M.O.J. 1988: *A Conceptual Approach to the Early Learning of Algebra Using a Computer*, unpublished Ph.D. Thesis, The University of Warwick.

Vinner S. & Hershkowitz R.1980 : 'Concept Images and some common cognitive paths in the development of some simple geometric concepts', *Proceedings of the Fourth International Conferenceof P.M.E.*, Berkeley,177-184.

Vinner S. 1983: 'Conflicts between definitions and intuitions: the case of the tangent', *Proceedings of P.M.E. 6*, Antwerp.