The Transition to Advanced Mathematical Thinking:

Functions, Limits, Infinity and Proof[1]

Introduction

Advanced mathematical thinking - as evidenced by publications in research journals - is characterized by two important components: *precise mathematical definitions* (including the statement of axioms in axiomatic theories) and *logical deductions* of theorems based upon them. However, the printed word is but the tip of the iceberg - the record of the final "precising phase", quite distinct from the creative phases of mathematical thinking in which inspirations and false turns play their part.

A major focus in mathematical education at the higher levels is to initiate the learner into the complete world of the professional mathematician, not only in terms of the rigour required, but also to provide the experience on which the concepts are founded. Traditionally this has been done through a gentle introduction to the mathematical concepts and the process of mathematical proof in school before progressing to present mathematics in a more formally organized and logical framework at college and university.

The move to more advanced mathematical thinking involves a difficult transition, from a position where concepts have an intuitive basis founded on experience, to one where they are specified by formal definitions and their properties re-constructed through logical deductions. During this transition (and long after) there will exist simultaneously in the mind earlier experiences and their properties, together with the growing body of deductive knowledge. Empirical research has shown that this produces a wide variety of cognitive conflict which can act as an obstacle to learning.

In this chapter we will look at the results of research into the conceptualization of several advanced concepts, including the notion of a function, limits and infinity and the process of mathematical proof, particularly during the transition phase from the later years of school to college and university. But first we must linger a little and consider the nature of our own perceptions of mathematical concepts, for even those of professional mathematicians contain idiosyncrasies dependent on personal experience.

Creases in the mind

> "The human mind", wrote Antoine Lavoisier, the French Chemist guillotined during the French Revolution, "gets creased into a way of seeing things." One might add that the evolving corporate mind suffers no less, since it perceives by indoctrination, from generation to generation.
>
> (Adrian Desmond, *The Hot-Blooded Dinosaurs*, page 128)

As we look back at the historical development of mathematics we see that successive generations develop their own corporate perception of mathematical ideas, based on mutual agreement over important concepts. The pre-Pythagorean Greeks believed that all numbers were rational, until the Pythagorean theorem revealed that the square root of two is not. Aristotelian dynamics suggested that the speed of a moving body is

---

[1] Published in Grouws D.A. (ed.) *Handbook of Research on MathematicsTeaching and Learning*, Macmillan, New York, 495–511.

proportional to the force applied until Newton's laws proposed that it is *acceleration* that is proportional to force, not speed. For two millenia Euclidean geometry was regarded as the pinnacle of deductive logic until nineteenth century mathematicians realized that there were theorems that depended on implicit assumptions (such as the fact that the diagonals of a rhombus lie *inside* the figure) which were not logical deductions from the axioms.

It would be a mistake to assume that at last we have "got it right" and that this generation is free of the internal conflicts and confusions of the past. On the contrary, we have our own share of corporate creases of the mind. (See, for example, Sierpiń ska 1985a, 1985b, 1987.) Many of the creases that we purport to see in students are actually present in ourselves and have been passed down in varyingly modified forms from generation to generation.

For example, the idea that a function y=f(x) is single-valued has become part of our mathematical culture and we may find it strange to see students asserting that a circle $x^2+y^2=1$ can be a function. Yet the term "implicit function" continues to be used in text-books to describe such an expression. I (to my eternal shame) find that I have published a computer program called the "implicit function plotter" which will draw, amongst other things, the graph of $x^2+y^2=1$. Likewise I find myself considering the draft of a new curriculum for the 16-19 age range in Britain which says of this equation: "strictly speaking, y is not a function of x because there is not a unique value of y for each value of x, but we might think of it as a 'double-valued' function from x to y". What are students to think? Can any of us, with hand on our heart state that we have never indulged in any vagaries of this kind? Let him (or her) who is without sin cast the first stone...

Concept definition and concept image

> What is a good definition? For the philosopher or the scientist, it is a definition which applies to all the objects to be defined, and applies only to them; it is that which satisfies the rules of logic. But in education it is not that; it is one that can be understood by the pupils. (Poincaré, 1908)

The "new mathematics" of the sixties was a valiant attempt to create an approach based on clear definitions of mathematical concepts, presented in a way that it was hoped that students would understand. But it failed to achieve all its high ideals. The problem is that the individual's method of thinking about mathematical concepts depends on more than just the form of words used in a definition:

> Within mathematical activity, mathematical notions are not only used according to their formal definition, but also through mental representations which may differ for different people. These 'individual models' are elaborated from 'spontaneous models' (models which pre-exist, before the learning of the mathematical notion and which originate, for example, in daily experience) interfering with the mathematical definition. We notice that the notion of limit denotes very often a bound you cannot cross over, which can, or cannot, be approached. It is sometimes viewed as reachable, sometimes as unreachable. (Cornu 1981)

Thus the experience of pupils prior to meeting formal definitions profoundly affects the way in which they form mental representations of those concepts. During the late seventies and early eighties many authors noted the mismatch between the concepts as formulated and conceived by formal mathematicians, and as interpreted by the student apprentice. For example, difficulties were noted in the understanding of the limiting process as secants tend to tangents (Orton 1977), the meaning of infinite decimals such as "nought point nine recurring" (Tall 1977), geometrical concepts (Vinner & Hershkowitz, 1980), the notion of function (Vinner 1983), limits and continuity (Tall & Vinner 1981, Sierpińska 1987), the meaning of the differential (Artigue 1986), convergence of sequences (Robert 1982), limits of functions (Ervynck 1983), the tangent (Vinner 1983, Tall 1987), infinite series (Davis 1982), infinite expressions (Borasi 1985), the intuition of infinity (Fischbein et al 1979), and so on.

To highlight the role played by the individual's conceptual structure, the terms "concept image" and "concept definition" were introduced in Vinner & Hershkowitz (1980) and later described 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 will refer to 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. (Tall & Vinner 1981, p.152)

On the other hand:

> The *concept definition* [is] a form of words used to specify that concept. (ibid.)

The consideration of conflicts in thinking is widespread in the literature:

> New knowledge often contradicts the old, and effective learning requires strategies to deal with such conflict. Sometimes the conflicting pieces of knowledge can be reconciled, sometimes one or the other must be abandoned, and sometimes the two can both be "kept around" if safely maintained in separate mental compartments. (Papert, 1980, page 121)

> In general, learning a new idea does not obliterate an earlier idea. When faced with a question or task the student now has *two* ideas, and may retrieve the new one or may retrieve the old one. What is at stake is not the possession or non-

> possession of the new idea; but rather the *selection* (often unconscious) of which one to retrieve. Combinations of the two ideas are also possible, often with strikingly nonsensical results. (Davis and Vinner,1986, page 284)

This is particularly applicable to the transition to advanced mathematical thinking when the mind simultaneously has concept images based on earlier experiences interacting with new ideas based on definitions and deductions. The very idea of *defining* a concept in a sentence, as opposed to *describing* it, is at first very difficult to comprehend - particularly when there are bound to be words in the definition which are not themselves defined. It is impossible to make a beginning without making some assumptions, and these are based upon the individual's concept image, not in any logically formulated concept definition!

Mathematical foundations and cognitive roots

> Burrow a while and build, broad on the roots of things.
> (Robert Browning, 1812-1889, *Abt Vogler*)

In building a curriculum it is natural to attempt to start from simple ideas and move steadily to more complex concepts as the student grows in experience. What better foundations to build upon than the definitions which have been evolved over many generations? The problem is that these definitions are both subtle and generative whilst the experiences of students are based on the evident and particular, with the result that the generative quality of the definitions is obscured by the students' specific concept image. For example, a function may be defined as a process which assigns to each element in one set (the domain) a unique element in another (the range). It is not possible to give the full range of possibilities embedded in this definition at the outset - that the sets involved may be sets of numbers, or of points in *n*-dimensional space, or geometrical shapes, or matrices, or any other type of object, including other functions - that the method of assignment might be through a formula, an iterative or recursive process, a geometrical transformation, a list of values, or any serendipity combination one desires, provided that it satisfies the criterion of assigning elements uniquely.

When students are first confronted with mathematical definitions it is almost inevitable that they will meet only a restricted range of possibilities that colours their concept images in a way that will cause future cognitive conflict.

Rather than deal initially with formal definitions which contain elements unfamiliar to the learner, it is preferable to attempt to find an approach which builds on concepts which have the dual role of being familiar to the students and also provide the basis for later mathematical development. Such a concept I term a *cognitive root*. These are not easy to find - they require a combination of empirical research (to find out what is appropriate to the student at the current stage of development) and mathematical knowledge (to be certain of the long-term mathematical relevance). A cognitive root is different from a mathematical foundation. Whilst a mathematical foundation is an appropriate starting point for a logical development of the subject, a cognitive root is more appropriate for curriculum development.

For example, the limit concept is a good example of a mathematical foundation - honed and made precise over the centuries by the combined efforts of many great mathematicians. But it proves to be difficult for students to use as a basis of their thinking and may not be a sound cognitive root for the beginning stages of the calculus. On the other hand, the idea that certain graphs look less curved as they are more highly magnified is intuitively appealing and can be discovered by any student playing with a graph plotter.The fact that this can grow into the formal theory of differential manifolds which are locally like *n*-dimensional space suggests that "local straightness" may prove to be a suitable cognitive root for the calculus. The case for local straightness is enhanced when it is realized that the solving of a (first-order) differential equation is essentially the reverse problem: to find a (locally straight) function which has a given gradient. It is possible, with software, to build a picture of an approximate solution enactively just by placing short line segments of the appropriate gradient end to end.

The function concept

> The keynote of Western culture is the function concept, a notion not even remotely hinted at by any earlier culture. And the function concept is anything but an extension or elaboration of previous number concepts – it is rather a complete emancipation from such notions. Schaaf (1930), p.500

The function concept, according to Kleiner (1989), "goes back 4000 years; 3700 of these consist of anticipations". Its evolution has led to a complex network of conceptions: the geometric image of a graph, the algebraic expression as a formula, the relationship between dependent and independent variables, an input-output machine allowing more general relationships, through to the modern set-theoretic definition (see, for example, Buck, 1970).

In the "New Math" there was a valiant attempt to build the function concept from a formal definition in terms of the cartesian product of sets A and B:

> Let A and B be sets, and let AxB denote the cartesian product of A and B. A subset f of AxB is a *function* if whenever (x_1,y_1) and (x_2,y_2) are elements of f and $x_1=x_2$, then $y_1=y_2$.

However, there is much empirical evidence to show that, though this definition is an excellent mathematical foundation, it may not be a good cognitive root. The "emancipation" from previous concepts suggested so eloquently by Schaaf over sixty years ago is mirrored in the total cognitive reconstruction which is necessary to use the new set-theoretic definition in place of earlier process-related notions. It is a reconstruction which students seem to find extremely difficult.

Malik (1980) highlighted the manner in which this definition represents a very different frame of thought from that experienced in traditional calculus emphasising the rule-based relationship between a dependent and independent variable.

Sierpiǹ ska focussed on the latter use of the function concept and asserted:

> The most fundamental conception of a function is that of a relationship between variable magnitudes. If this is not developed, representations such as equations and graphs lose their meaning and become isolated from one another... Introducing functions to young students by their elaborate modern definition is a didactical error - an antididactical inversion. (Sierpiǹ ska, 1988, p. 572)

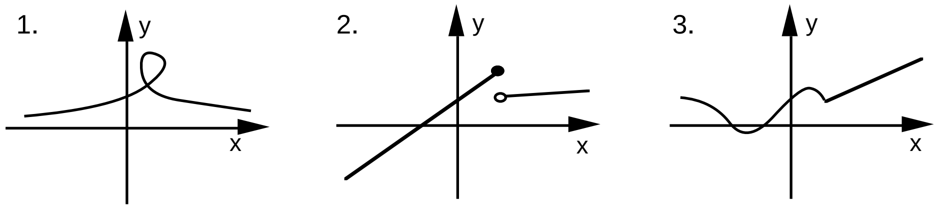
Empirical research shows that, even when students are given such a formal definition, their overwhelming experience from examples of functions with implicit common properties causes them to develop a personal concept image of a function which implicitly has these properties. For instance, if the functions encountered are given mainly in terms of formulae, this causes many students to believe that the existence of a formula is essential for a function.

Dreyfus & Vinner (1982, 1989) asked[1] a cross-section of 271 college students and 36 teachers a number of conceptual questions about functions (figure 1).

The responses to the notion of function (question 5) included not only the standard definition (each value of x corresponds to precisely one value of y), but also variants such as:

[1]The original questions were in Hebrew.

Does there exist a function whose graph is:



4. Does there exist a function which assigns to every number different from zero its square and to 0 it assigns 1?

5. What in your opinion is a function?

Figure 1 : What do students think about functions?

a correspondence between two variables
a rule of correspondence
a manipulation or operation (on one number to obtain another)
a formula, algebraic term, or equation
a graph, y=f(x), etc.

However, the responses to the first four questions were not always in accord with these notions. Table 1 shows the percentage of students whose responses were adjudged correct:

Mathematical Level:	Low	Intermediate	High	Math Majors	Teachers
Question: 1	55%	66%	64%	74%	97%
2	27%	48%	67%	86%	94%
3	36%	40%	53%	72%	94%
4	9%	22%	50%	60%	75%

Table 1 : Student Responses to function questions

The percentages improve with ability and experience, but non-mathematics majors in particular have a high percentage of incorrect responses.

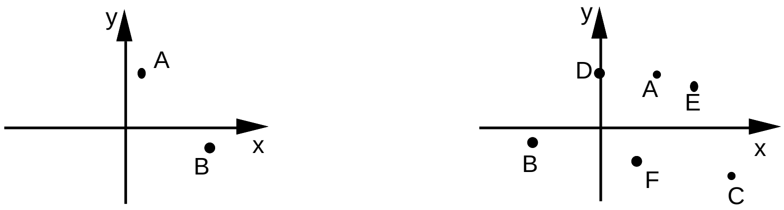
The reasons for the responses include not only the standard definition and the variants above, but also evoked concept images such as:

The graph is “continuous” or changes its character (e.g. two different straight lines),
the domain of the function “splits”,
there is an exceptional point.

Although the original questions are somewhat out of the ordinary, similar results have been replicated in other studies (e.g. Vinner 1983, Barnes 1988, Markovits, Eylon & Bruckheimer 1986, 1988).

Markovits *et al* (1986, 1988) conclude that the complexity of the modern definition causes problems because of the number of different components (domain, range and rule), yet little emphasis is placed on domain and range at school level, resulting in stress being placed on the rule or relationship (which is usually given as a formula). Early emphasis on straight line graphs seems to cause students to evoke linear graphs when asked to consider possible functions through give points (figure 2).

In the given coordinate system, draw the graph of a function such that the coordinates of each of the points A,B, [C, D, E, F] represent a pre-image and the corresponding image of the function:



The number of different such functions that can be drawn is -

- 0
- 1
- 2
- more than 2 but fewer than 10
- more than 10 but not infinite
- infinite.

Explain your answer.

Figure 2 : More function questions

The first figure often evoked a straight line allowing only one function because “two points can be connected by only one straight line”. The second caused problems, perhaps because of the disposition of the points seemingly on two different lines : “If I draw a function such that all the points are on it, what will happen is for every *x* there will be two *y*, and it will not be a function”.

The authors observe (page 54):

> Their conception of functions as linear would seem to be influenced by geometry (which they learn simultaneously with algebra) and also by the time spent in the curriculum exclusively on linear functions.

Barnes (1988) asked questions of grade 11 school students and university students about different representations, for instance, whether expressions such as

$y=4$,

$x^2+y^2=1$,

$$y=\begin{cases} 0 & \text{if } x\leq 0 \\ x & \text{if } 0<x\leq 1 \\ 2-x & \text{if } x>1 \end{cases}$$

define *y* as a function of *x*. A majority decided that the first did not, because the value of *y* does not depend on *x*, many decided that the second *is* a function (because it is a circle, which is familiar to them), whilst the third presented difficulties because it appeared to define not *one* function but *several*.

When asked which graphs represented *y* as a function of *x*, including those in figure 3,

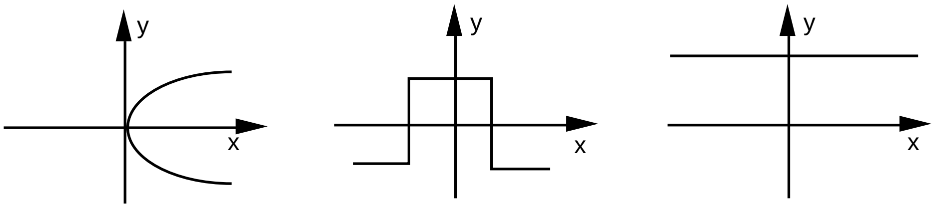


Figure 3 : Are these graphs of functions?

students responded in a variety of ways. The first graph evoked images such as "It's more like *x* is a function of *y*" or "It's a rotated function" or "That's $y=x^2$ so it's a function". The second was almost universally regarded as not being a function, not because it has vertical line segments, but because "it looks strange", or "it's not smooth and continuous", or "It's too hard to define it". The last one, in contrast to the algebraic expression $y=4$ was regarded as being a function by *all* the university students, though some of the school students were concerned that *y* was always the same. Some of the university students, who saw the horizontal line as a function, had asserted earlier that y=*4* was not a function, but now realized there was a conflict. Some, but not all, wanted to go back to the earlier question to modify their response.

At this stage, it would be of interest for the reader to look back at some of these questions to see the creases in the mind that we all share. For example, the questions related to figure 1 assume that *y* is being considered as a (possible) function of *x*. The first picture could so easily be what is often described as a "parametric graph" – the image of a function from an interval to the plane. The cartoon-like blobs in the second picture are a convention to represent a discontinuity; if you think about it, you will realize that it does not truly represent the ordered pairs on the graph in the neighbourhood of the discontinuity. In fact a physical graph is only a rough representation of a function, with subtle conceptual difficulties, such as the fact that younger children see the graph as a curve and not as a set of points (Kerslake 1977).

Could the middle graph of figure 3 represent a function? It seems not, yet it *could* do if the "vertical lines" were actually very steep but not vertical, say in the form:

$y=-1 \quad \text{if } |x|\geq 1+k,$

$y=\dfrac{1-|x|}{k} \quad \text{if } 1-k<|x|<1+k,$

$y=1 \quad \text{if } |x|\leq 1-k$

where *k* is very small (say $k=1/1000$).

Few students would be aware of such possibilities. However they indicate the implicit creases in our minds that present students with a minefield through which we trust they will choose a consistent path. Is it any wonder that so many fail?

Even greater difficulties with the function concept are encountered with the variety of different representations (graph, arrow diagram, formula, table, verbal description, etc) and the relationships between them (Thomas 1975, Dorofeev 1978, Dreyfus & Eisenberg 1982, Janvier 1987). For instance, Dreyfus and Eisenberg (1987) found that students have considerable difficulties relating the algebra of transformations (such shifts $f(x) \to f(x)+k$, $f(x) \to f(x+k)$ and stretches $f(x) \to kf(x)$, $f(x) \to f(kx)$) to their corresponding graphical representations. Of these, the transformations in the domain $f(x) \to f(x+k)$, $f(x) \to f(kx)$ naturally proved to be the more difficult.

Even (1988) studied the concept of function in prospective mathematics teachers. She found similar difficulties with student teachers in the final year of their mathematics studies.

> ... Many of them ignored the arbitrary nature of the relationship between the two sets on which the function is defined ... Some expected functions to always be representable by an expression. Others expected all functions to be continuous. Still others accepted only "reasonable" graphs, etc. (Even, 1988 page 216)

> ... Can we expect teachers to be able to teach according to a modern definition of function, as it now appears in modern texts, while their conception of function is more restricted, more primitive? The participants' incomplete conception of function is problematic and may contribute to the cycle of discrepancies between concept definition and concept image of functions in students ... keeping the students' concept image of functions similar to the one from the 18th century. (ibid page 217-8)

Given the creases in the minds of prospective teachers, is it any wonder that it proves continually difficult to address the deep problems with the function concept in students?

In recent years the computer has been harnessed to introduce the function concept. Many of the initial moves have focussed on the graphical representation of functions (Demana & Waits 1988, Dugdale, 1982, Goldenberg et al 1988, Schwartz 1990, Yerushalmy, 1990, 1991). These techniques change the conception of a function from a rule-based pointwise process to a global visualization of overall behaviour. Entirely new approaches are possible, for example, to view the qualitative shape of graphs to suggest algebraic or trigonometric relationships (Dugdale, 1989, Schwartz 1990).

This brings on the one hand a great increase in potential power and, on the other, greater potential for misinterpretations of the graphical representation. For instance, the graph may look very different when drawn over different ranges (Demana & Waits 1988) and there may be visual illusions created by the changing scales of each axis (Goldenberg et al 1988). The technology places enormous power in the hands of students but serious research is necessary (and currently in progress) to gain insights into student conceptions generated by its use.

For example, most of the graph-plotting software initially available on microcomputers only accepted functions given by formulae, implicitly reinforcing the student's restricted concept image of a function as a formula. An exception is ANUGraph from the Australian National University, which allows functions to be defined by different formulae on several domains.

Only recently have graph-drawing programs appeared that allow the function notation $f(x)$. For instance, the School Mathematics Project "Function Analyser" in Britain, allows functions to be typed in terms of expressions such as $g(t)=t+\sin t$, $f(u)=e^u$, and these in turn may be used in expressions to draw graphs such as $y=g(x)+1$, $y=g(x+1)$ or $y=f(g(x))$. But this is still limited to functions given by formulae.

The "Triple Representation Model" (Schwarz & Bruckheimer 1988), offers facilities to draw the graphs of functions, calculate and plot numerical values of functions, and step-search over an interval to find points which satisfy a specified equality or inequality. Here the functions are sums, differences, products and compositions of rational, absolute value, square root and integer part functions, defined on continuous or discrete domains. The software can be used for problem-solving activities, revealing, for example, how students use different representations to find solutions of equalities and inequalities, by "zooming in" on points where graphs cross, or using a combination of numerical evaluation and step-search strategies. Schwarz *et al* (1988) report that the software "enables students to reach higher cognitive levels in functional reasoning". For instance, the experiences with more general graphs significantly diminishes the "linear graph" response to questions asking for a variety of function graphs through a given set of points. The use of the software leads to more sophisticated strategies for solving equations using the facilities provided.

There is a veritable explosion of the use of graphic calculators and graphical software on computers for the drawing of graphs of functions, some of which is indicated in a broad-ranging review of the relationships between functions and graphs by Leinhardt et al (1990). Indeed, this is an excellent source of information and references for further study of empirical research into the function concept in general, indicating the depth of complexity and difficulty of the topic.

In emphasizing the many representations of the function concept: formula, graph, variable relationship and so on, the central idea of function as a *process* is often overlooked. For example, although graphs are often represented as an excellent way to think of a function, very few students seem to relate the graph to the underlying functional process (take a point on the x-axis, trace a vertical line to the graph and then a horizontal line to the y-axis to find the value of $y=\mathrm{f}(x)$). Instead students see a graph simply as an object: a static curve (Dubinsky 1990).

To consider the concept of function *as a process*, Dubinsky and his co-workers introduced students to the function concept via programming Ayers *et al* (1988). Using the Unix operating system a number of commands were prepared for student use, some operating on numbers and some on text. The intention was to help the students think of a function both dynamically as a process and encapsulate it statically as a mental object on which operations such as function composition may be performed. Though the number of students involved was small, there was evidence to support the hypothesis that the computer experiences were more effective for the experimental students than traditional paper and pencil exercises carried out by a control group.

From this experiment, Dubinsky progressed to the idea of programming the more general notion of function on finite sets using the language ISETL (Schwartz *et al* 1986). This programming language allows students to handle functions as arbitrary sets of ordered pairs as well as procedures, and to construct operations such as function composition in a mathematical way. More recent implementations of ISETL allow the user to graph the functions so constructed.

Empirical research shows that students can learn to think of a function as a process by programming a procedure on the computer to carry out the process (Breidenbach, *et al* to appear). At a later stage, a function defined in this way can be used as an input to another procedure, hence encapsulating the process *as an object*. This suggests that the act of programming function procedures may provide a cognitive root from which the formal concept may grow. The ISETL language also provides a programming environment in which the learner may reflect on the difficult transition from function as process to function as object.

The research discussed in this section shows a wide variety of approaches to the complexity of the function concept. Some gain can be made in improving understanding and problem-solving abilities in specific areas of the function concept, but there is no appears to be no universal panacea. The idea of function as a process may prove to be a suitable cognitive root for the formal concept, but along the line of cognitive development there are obstacles to be overcome, including the encapsulation of the process as a single concept and the relating of this concept to its many and varied alternative representations. It remains a large and complex schema of ideas requiring a broad range of experience to grasp in any generality.

The notion of a limit

> Est modus in rebus, sunt certi denique fines,
> Quos Ultra citraque nequit consistere rectum,
> (Things have their due measure; there are ultimately fixed limits,
> beyond which, or short of which, something must be wrong.)
>
> (Horace 65-8 B.C. *Satires*)

Although the function concept is central to modern mathematics, it is the concept of a limit which signifies a move to a higher plane of mathematical thinking. As Cornu observed (1983), this is the first mathematical concept that students meet where one does not find the result by a straightforward mathematical computation. Instead it is "surrounded with mystery", in which "one must arrive at one's destination by a circuitous route".

Limits occur in many different mathematical contexts, including the limit of a sequence, of a series, of a function (of f(x) as $x \to a$, or as $x \to \infty$), in the notion of continuity, of differentiability, of integration. In a mathematical sense it would be appropriate to distinguish between these various different types of limit, for example, the discrete limit of a sequence (a_n) as $n \to \infty$ and the continuous limit of f(x) as $x \to a$. However, empirical research shows common difficulties for beginners across the various mathematical categories.

For example, the word "limit" itself has many connotations in everyday life which are at variance with the mathematical idea. An everyday limit is often something which cannot or should not to be passed, such as a "speed limit". The terminology associated with mathematical limiting processes includes phrases such as "tends to", "approaches", or "gets close to" which again have colloquial meanings differing from the mathematical meanings. For instance when these phrases are used in relation to a sequence *approaching* a limit, they invariably carry the implication that the terms of the sequence *cannot equal* the limit (Schwarzenberger & Tall, 1977).

The problem of handling limits is exacerbated by restricted concept images of sequences and functions, for example students are often introduced to the notion of a sequence where the terms are given as a formula. If one wished to show that some terms of a sequence might equal the limit , one might try to consider the sequence 1,0, $\frac{1}{2}$, 0, $\frac{1}{3}$, 0, ... but students who view the terms of a sequence as a formula may insist that this is not *one* sequence, but *two*; the odd terms form a harmonic sequence 1, $\frac{1}{2}$, $\frac{1}{3}$, ... which tends down to zero and the even terms are constants, which *are* zero (Tall 1980b).

Davis and Vinner (1986) suggest that there are seemingly *unavoidable* misconception stages with the notion of a limit. One is the influence of language, mentioned earlier, in which the terms remind us of ideas that intrude into the mathematics. In addition to the *words* , there are the *ideas* that these words conjure up, which have their origins in earlier experiences. Although the authors attempted to teach a course in which the word "limit" was not used in the initial stages, they eventually concluded that "avoiding appeals to such pre-mathematical mental representation fragments may very well be futile". Another source of misconceptions is the sheer *complexity* of the ideas, which cannot appear "instantaneously in complete and mature form", so that "some parts of the idea will get adequate representations before other parts will". *Specific examples* are likely to dominate the learning, for instance they found that monotonic sequences dominated their early examples, so it was not surprising that they dominated the student's concept images. This could lead to a *misinterpretation of one's own* *experience*, for instance the fact that students dealt with many examples of sequences whose terms were given by a formula caused them to mistakenly assume that a simple algebraic formula for the *n*th term a_n is an essential part of the theory.

Most of the informal ideas of limit carry with them a dynamic feeling of something approaching the limiting value, for instance, as n increases, the sum

$$1 + \frac{1}{2} + \ldots + \frac{1}{2^n}$$

approaches the limit 2. This has an inevitable cognitive consequence which I term the "generic limit property" (Tall 1986): the belief that any property common to all terms of a sequence also holds of the limit. It is a belief with worthy historical precursors, for example in Leibniz's "Principle of continuity" (stated in a letter to Bayle), that

> In any supposed transition, ending in any terminus, it is permissible to institute a general reasoning, in which the final terminus may also be included.

It permeates the history of mathematics, for instance, in Cauchy's belief that the limit of continuous functions must again be continuous. And it remains as a crease in the mind of today, in such ideas that the limit of the sequence

0.9, 0.99, 0.999, ...

must be *less* than one - because all the terms are less than one. Thus "nought point nine recurring" is "just less than one". (see, for example, Schwarzenberger & Tall, 1977, Tall & Vinner, 1981, Cornu 1983 etc). Cornu (1983) studied this in greater detail and found a whole array of beliefs, for instance that "0.9, 0.99, ... *tends* to nought point nine recurring, but has *limit* one" (because it "tends" to have the property of 0.99999..., but cannot pass the "limit" one).

It will come as no surprise that attempting to "simplify" the limit notion by using everyday language can lead to serious conceptual problems. Orton (1980) investigated students concepts of limit using a "staircase with treads" where extra half-size treads are inserted between each tread, then the process repeated successively with treads half this size again. (Figure 4.)

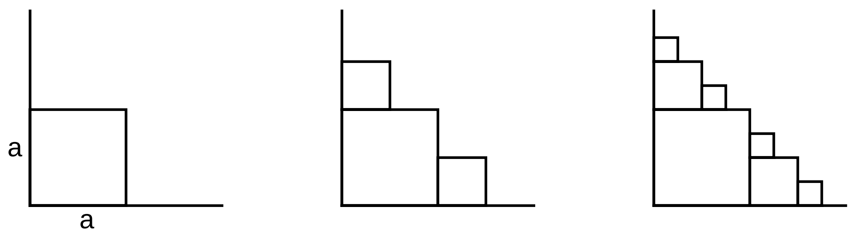


Figure 4 : A limiting staircase

In an interview he posed the questions:

> (a) If this procedure is repeated indefinitely, what is the final result?
>
> (b) How many times will extra steps have to be placed before this "final result" is reached?
>
> (c) What is the area of the final shape in terms of "*a*", i.e. what is the area below the "final staircase"?

If a formula was given in (c) he asked:

> Can you use this formula to obtain the 'final term' or limit of the sequence ?

He justified the use of this terminology by stating:

> The expression "final term" was again used in an attempt to help the students understand the meaning of limits.

He is surely not alone in his attempt to help the student by an informal presentation. But a phrase such as "the final staircase" is likely to create a generic limit concept in which the student imagines a staircase with an "infinite number of steps", and this is precisely the response that it evoked.

Faced with such difficulties in the dynamic notion of a limit, it will come as no surprise that the formal definition is also fraught with cognitive problems. Even the phrase "given an epsilon greater than zero... " may be interpreted as taking epsilon to be "arbitrarily small" and this in turn can lead to the symbol generically representing an "arbitrarily small" number:

> Everything occurs as if there exist very small numbers, smaller than "real" numbers, but nevertheless not zero. The symbol ε represents for many students a symbol of this type: ε is smaller than all real numbers, but not zero. (Cornu 1983)

In the same way, in the calculus, the introduction of symbols δx (used in the UK for a small finite increment in x) and dx (as part of the dy/dx notation) leads generically to the idea that there exist numbers that are arbitrarily, or "infinitesimally", small (Orton 1980, Tall 1980, Cornu 1983).

The introduction of the formal notion of limit does not obliterate more primitive dynamic notions, indeed, we often continue to nurture dynamic imagery in our teaching to give an intuitive flavour to rigorous proofs.

Robert (1982) studied the notion of limit of a sequence as perceived by 1380 students at various levels in school and university. She asked how the students might explain the notion of a convergent sequence to a pupil of 14 or 15 years old (a question that is more likely to evoke a concept image than the formal definition). She classified the responses into four main categories:

> 1. *Monotonic & Dynamic Monotonic* (12%)
> "a convergent sequence is an increasing sequence bounded above (or decreasing bounded below)
> "a convergent sequence is an increasing (or decreasing) sequence which approaches a limit"
>
> 2. *Dynamic* (35%)
> "u_n tends to l ", "u_n approaches l ", "the distance from u_n to l becomes small"
> "The values approach a number more and more closely"
>
> 3. *Static* (13%)
> "The u_n are in an interval near l ", "The u_n are grouped round l ",
> "u_n is as close as you like to l "
>
> 4. *Mixed* (14%)
> A mixture of those above.

In addition, 4% gave the formal definition, 5% did not attempt the question and the remainder gave incomplete or false statements, such as "u_n doesn't go past l " or "u_n stays below l ".

The fact that a student evokes a particular image does not mean the absence of other images:

> The presentation by a student of an old (and incorrect) idea cannot be taken as evidence that the student does NOT know the correct idea. In many cases the student knows both but has retrieved the old idea. (Davis & Vinner 1986, p.284)

In particular, Robert's request for an explanation suitable for a 14 or 15 year-old seems to exclude the formal definition because of its difficulty. This difficulty is confirmed by Tall & Vinner (1981) who asked seventy highly qualified first year university mathematics students to write down a definition of $\lim_{x \to a} f(x) = c$ (if they knew one). They had just arrived from school and would be expected to have been given a dynamic definition (f(x) gets close to c as x gets close to a), though some might have been shown the formal definition. Those who replied did so as follows:

	correct	incorrect
formal	4	14
dynamic	27	4

Table 2 : Student definitions of the limit concept

Thus the majority of those who recalled the (easier) dynamic definition could state it correctly, whilst the majority of those who chose to give the formal definition were not able to recall it in a satisfactory way, mis-stating it in various ways such as:

$|f(x)-c| < \varepsilon$ for all positive values of ε with x sufficiently close to a

As $x \to a$, $c-\varepsilon \le f(x) \le c+\varepsilon$ for all $n > n_0$

$|f(n)-f(n+1)| < \varepsilon$ for all $n >$ given N_0.

Teaching the notion of limit using the computer has, on the whole, fared badly. Regular computing languages, such as BASIC, Pascal or C, hold numbers in a fixed number of memory locations which can lead to serious problems of accuracy when calculating a limit such as

$$\lim_{h \to 0} \frac{\sin(x+h)-\sin x}{h}$$

When h is small both numerator and denominator are tiny numbers whose quotient is likely to be highly inaccurate. For instance, for $x=\pi/3$, the limit as h tends to 0 should be $\frac{1}{2}$, but on a typical micro, taking $h=1/10^n$ for $n=1$ to 10 gives the sequence:

```
0.455901884
0.495661539
0.499954913
0.499980524
0.499654561
0.500585884
0.465661287
0.232830644
0
```

which hardly gets close to 0.5.

Numerical ideas of limits in such contexts must therefore be combined with discussion of accuracy of computer arithmetic.

Symbolic treatments of limits do not always fare better. The expression

```
((x+h)^2-x^2)/h
```

typed into Derive (Rich et al 1989) is prettily printed as

```
          2  2
(x+h) -x
---------
    h
```

and may be automatically simplified to

```
2x+h
```

but there is no warning about the case h=0.

The more general expression

```
     n  n
(x+h) -x
---------
    h
```

only simplifies to

```
     n    n
(x+h)    x
----- - --- .
  h      h
```

Derive’s limit option applied to this expression, as h tends to 0, gives not nx^{n-1}, but

```
 n LN(x)-LN(x/n)
ê
```

which may be suitable for sophisticated investigation, but is hardly appropriate for a beginner[1]. It illustrates the difficulties encountered when one tries to program symbol manipulation. It is hard enough to do, but far harder to get the expression into a form that may be desired.

All the evidence points once more to the fact that, although the limit concept (in a formal sense) is a good mathematical foundation, it fails to be an appropriate cognitive root. If it is difficult to start with the limit process in subjects such as the calculus, what alternatives are available? Instead of introducing *explicit* limit ideas in differentiation, Tall (1986) begins by magnifying graphs. This builds on the thesis that a cognitive root for the calculus is the idea that a differentiable function has a graph which magnifies to “look straight”. To give a rich concept image, the software (Tall, Blokland & Kok 1990) includes not only standard functions, but also functions which are so wrinkled, that no matter how highly magnified, they *never* look straight. Thus *in the very first lesson* in the calculus it is possible to explore functions which are locally straight everywhere, functions which have different left and right gradients at certain points (because the graph magnifies to reveal a corner) and functions that are locally straight nowhere (because they are too wrinkled). This allows students to build a much richer concept image including examples of differentiable functions, functions having different left and right derivatives, and non-differentiable functions – providing cognitive roots on which formal theories may later be grafted.

It is not an easy path. But this is true of life itself. There is no royal road, as Euclid is said to have remarked to Ptolemy. Given the complexity of the limit concept, the road ahead is surely not an attempt to ease the student’s path by attempting to avoid difficulties, for the over-simplification produces inappropriate concept images which only store up problems for later. A more helpful route is to provide the rich experience

1 In Derive version 2.0 this has been changed to give nx^{n-1}.

necessary to enable the student to attempt to to confront the difficulties and negotiate a more stable concept, mindful of the possible pitfalls.

Thinking about infinity

> To see a World in a Grain of Sand,
> And Heaven in a wild flower,
> Hold Infinity in the palm of your hand
> And Eternity in an hour.
>
> William Blake 1757-1827 (*Auguries of Innocence*)

Thoughts of infinity touch us all at some time or other as we contemplate the puny nature of our finite existence in the vastness of the universe. Research into the nature of students' concepts of infinity is probably more clouded by the creases in the minds of the researcher than any other. For what do we mean by "infinity"? It would be useful for the reader to pause a moment and think what infinity means to him or her before reading on.

Historically, philosophers distinguished between *actual* infinity ("there are an infinite number of whole numbers") and *potential* infinity ("for any whole number there is always one bigger"). In modern times actual infinity is interpreted using Cantor's theory of cardinal numbers in terms of one-one correspondences between sets. An *infinite* set is one which can be put in one-one correspondence with a proper subset. Thus the natural numbers $\{1,2,3,\ldots, n, \ldots\}$ form an infinite set because they can be put into one-one correspondence with the even numbers $\{2,4,6,\ldots,2n, \ldots\}$ in which n corresponds to $2n$. It is this cardinal form of infinity which is prominent in modern mathematics.

But there are properties of cardinal infinity which many find difficult to come to terms with, for instance that a set can have "as many" elements as a proper subset. In cardinal number terms there are as many natural numbers as rationals, as many points on a unit real line segment as on a real line segment length two, or on a real line as in a square, yet there are far more real numbers than rationals. Where is the consistency?

A number of research studies are based on the inconsistency between the *cardinal infinity* of Cantor and our *intuitions*. Here there are creases in our minds born of our experiences comparing infinite sets which children, with their different experiences, may not share. Such research on infinity is likely to say as much about the nature of our own conceptions as it does about the conflicts in the minds of children. For this reason it is essential that we briefly consider the nature of various conceptions of infinity before we proceed.

In Tall (1981) I suggested that experiences of infinity that children encounter rarely relate to the action of comparing sets, which means that they rarely lay the cognitive roots for the cardinal concept of infinity. For instance, when a child thinks of a "point on a line", it may be in the manner of a pencil mark (or it may be something entirely different, for example the "point" on a sword). A pencil mark has finite size. A child who views a point has having a tiny finite size is likely to develop a generic concept of a point which has an extremely small size.

If a line segment is made up of such points, then there will be a finite number, say a hundred, points to make it up. A line segment of twice the length will require twice the number, say two hundred. The only way that the double length line segment has the same number of points is if the points are twice the size! In extrapolating these ideas to the infinite case, a natural generic concept would be to have a kind of infinity with an infinite number of infinitesimally small points in a unit segment and twice as many in a segment twice the length. I once suggested this to a mathematical colleague who laughed at the naïvety of it all and said, yes, there were twice as many reals in a line of twice the length - one had $\aleph_0$ elements in it, the other had $2\aleph_0$ elements, and by cardinal arithmetic, $\aleph_0 = 2\aleph_0$! He thought it was extremely naïve to think otherwise.

Let us imagine N intervals in a unit length, each of length $1/N$. If N is very large, $1/N$ is very small. A line twice the length will have $2N$ intervals of the same size, where $2N$ is even larger, and certainly not equal to N. Non-standard analysis allows us to let N to be an element in an ordered field bigger than the real numbers, so that N is larger (in the given order) than any real number. In this technical sense, N is "infinite". It follows by manipulating the order relation that $1/N$ is smaller than any positive real and so, in this technical sense, it is "infinitesimal" (see Tall 1980a, 1981). Thus non-standard analysis allows a line to be made up of an infinite number of tiny line segments of infinitesimal size. A line of twice the length will have twice the number of points of the same size. Unlike cardinal infinity, the non-standard infinite numbers $2N$ and N are not equal - one is bigger than the other - just as in the intuition of a child. Thus, although the child's concept of infinity conflicts with cardinal infinity, it has properties which are consonant with non-standard infinity.

As we consider the concept of infinity during the transition to advanced mathematical thinking, we now become aware of wider possibilities. *There is more than one notion of infinity*. The symbol ∞, used in phrases such as "the limit as n tends to ∞", represents the idea of *potential* infinity. Students are usually told not to think of it as a genuine number, yet they may be confused to find it verbalized in many contexts as if it were. In addition there are at least three notions of "actual infinity": *cardinal infinity* (extending the notion of counting via the comparison of sets – the favoured form of infinity by mathematicians), *ordinal infinity* (the concept proposed by Cantor in terms of comparison of *ordered* sets), and the notion of non-standard infinity (generalizing the notion of *measuring* from real numbers to a larger ordered field). For simplicity I term this non-standard infinity *measuring infinity.* All these kinds of infinity are logical entities appropriate for study in advanced mathematics. In judging the intuitions of a child we should not make the mistake of considering only one kind of infinity - cardinal infinity - as the only true mathematical notion.

For example, the "infinite staircase" response to the question at the end of the last section is a perfectly reasonable non-standard response, though it is rejected by standard analysis. Similarly, the idea that 0.999.. to an infinite number of places, say N, is infinitesimally smaller than 1 (by the infinitesimal quantity $1/10^N$).

Tall (1980b) asked students to compute various limits, including the limits of

$$\frac{n^2}{n^2+1} \text{ and } \frac{n^5}{(1.1)^n}$$

as n tends to infinity. A student who wrote

$$\frac{n^2}{n^2+1} \to \frac{\infty}{\infty} = 1$$

was shown that a similar argument would give

$$\frac{n^5}{(1.1)^n} \to \frac{\infty}{\infty} = 1$$

but replied firmly

> "no it wouldn't, because in this case the denominator is a *bigger* infinity, and the result would be zero".

This sense of infinities of different sizes is not a cardinal concept - it is an extrapolation of experiences in arithmetic closer to *measuring* infinity, than *cardinal* infinity.

Fischbein et al (1979) give another clear example of measuring infinity where

$$1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\ldots$$

is stated to be

> $s = 2 - \frac{1}{\infty}$, because there is no end to the sum of segments.

Here it is the potential infinity of the limiting process that leads to a generic concept of measuring infinity. The suggested limit is typical of all the terms: just less than 2. The arithmetic fits nicely with non-standard analysis, but not with cardinal numbers where infinities cannot be divided.

> Most experiences with limits relate to things getting *large*, or *small*, or *close* to one another. All of these extrapolate experience from *arithmetic* rather than comparisons between sets and are more likely to evoke *measuring infinity*, rather that *cardinal infinity*. It follows that the ideas of *limits* and *infinity*, which are often considered together, *relate to two different and conflicting paradigms.*

Many different ideas of infinity can occur in different students in a given class. Sierpińska (1987) analysed the concept images of 31 sixteen year-old pre-calculus mathematics and physics students, and classified the students into groups which she labelled with a single name for each group:

> Michael and Christopher are *unconscious infinitists* (at least at the beginning): they say "infinite", but think "very big". ... For both of them the limit should be the last value of the term ... for Michael this last value is either plus infinity (a very big positive number) or minus infinity. ... It is not so for Christopher who is more receptive to the dynamic changes of values of the terms. The last value is not always tending to infinity, it may tend to some small and known number.
>
> George is a *conscious infinitist*: Infinity is about something metaphysical, difficult to grasp with precise definitions. If mathematics is to be an exact science then one should avoid speaking about infinity and speak about finite numbers only. In formulating general laws one can use letters denoting concrete but arbitrary finite numbers. In describing the behaviour of sequences *the most important thing is to characterize the nth term* by writing the general formula. For a given *n* one can then compute the exact value of the term or one can give an approximation of this value.
>
> Paul and Robert are *kinetic infinitists*: the idea of infinity in them is connected with the idea of time. ... Paul is a *potentialist*: To think of some whole, a set or a sequence, one has to run in thought through every element of it. It is impossible to think this way of an infinite number of elements. The construction of an infinite set or sequence can never be completed. Infinity exists potentially only. Robert is a *potential actualist*: it is possible [for him] to make a "jump to infinity" in thought: the infinity can potentially be ultimately actualized. For both, Paul and Robert, the important thing is to see how the terms of the sequence change, if there is a tendency to approach some fixed value. For Paul, even if the terms of a sequence come closer and closer so as to differ less than any given value they will *never reach it*. Robert thinks *theoretically* the terms will *reach it in the infinity.*

Fischbein (*et al*) (1978, 1979, 1981) investigated a number of conflicts inherent between the different conceptions of infinity, for instance, the conflict between the intuition of the single potential infinity and the many infinities of cardinal number theory, or the conflict between the finite number of points that may be marked physically on a line compared with the infinite number of points that are theoretically possible. He distinguishes between "primary intuitions" which are our common heritage and "secondary intuitions" which come from more specialized experiences. Thus the idea of potential infinity is a primary intuition, but it takes considerable experience of cardinal infinity to develop appropriate secondary intuitions as these conflict with deeply held convictions (such as "the whole is greater than the part").

Tirosh (1985) continued the work of Fischbein et al by designing a teaching program on "finite and infinite sets" for grade 10 students, taking their intuitive background into account (for example, the fact that they might appeal to the "part-whole" principle to declare that a set was bigger than a proper subset). The pupils were presented with quotations from mathematicians on the puzzling aspects of infinite sets, to encourage them to feel that it was legitimate to face such conflicts. It was found that it was possible to improve their understanding of the Cantorian theory by using dynamic teaching methods, including the open discussion of intuitive conflicts.

Other research has addressed the alternative paradigm of non-standard analysis. Sullivan (1976) studied the effectiveness of teaching the calculus at college level from a non-standard viewpoint which combined axioms for the real numbers and a larger set of hyperreal numbers containing infinite and infinitesimal elements (Keisler 1976). The approach is given a strong geometric flavour using a pictorial interpretation of these elements using "microscopes" and "telescopes". She found that the students following the experimental course scored at least as well as a control group in regular analysis problems (ε-δ definitions, calculating limits, proofs, and applications) but were better at aspects of the course which had alternative interpretations using infinitesimal arguments. The latter tend to seem easier, partly because they do not involve as many quantifiers as the standard definitions, and partly because they extend intuitive experiences of "getting small" in the limit process.

Despite this empirical proof for the success of an approach using infinitesimals, the approach to calculus in higher education has hardly changed. There are genuine reasons for this, including the intrinsic sophistication of the non-standard ideas which depend on logic of the depth of the axiom of choice. But there are also prejudices arising from traditional mathematical analysis and its links with the theory of Cantor. The creases of the mind run deep.

It is important to complement the study of student difficulties with possible sources of difficulty in the mind of the teacher. Evidence from pre-service elementary teachers enrolled in an upper-division course in mathematics methods at a large university revealed widespread inconsistencies (Wheeler and Martin, 1987, 1988). Questions asking for explanations of the symbol ∞ and the final three dots in the expression "1, 5, 25, 125, 625, ..." showed that more than half the subjects were unfamiliar with the symbolism. Responses to "what is infinity?" referred either to an unending process - "the numbers go on without stopping", or to a recursive process - "no matter what number you say, there is always one greater simply by adding one to it". In either case the predominant notion of infinity evoked is *potential infinity*.

The question:

> TRUE or FALSE: Every line segment contains an infinite number of points,

(which could evoke either potential, cardinal or measuring infinity) had 39 true responses, 24 false, and 7 without a reply, whilst:

> TRUE or FALSE: There exists a smallest fraction greater than zero,

yielded 28 true, 29 false, and 13 not responding.

When cross-referenced the responses revealed the great majority of students “holding incomplete and inconsistent concepts of infinity” and individual written responses showed a wide variety of evoked concept images riddled with conflicts and inconsistencies. However, it is also interesting to ask whether the concept of infinity provoked by asking the meaning of “...” (potential infinity) is the same kind of infinity as the number of points in a line segment (cardinal infinity). In order to research into the beliefs held by students and to classify those beliefs, it is important first to analyse the concepts concerned and the kind of concept images generated by various experiences without imbuing them with a classical mathematical prejudice.

Mathematical proof

> For nothing worthy proving can be proven
> Nor yet disproven: wherefore thou be wise.
> Cleave ever to the sunnier side of doubt.
> (Alfred Lord Tennyson, 1809-1892, *The Ancient Sage*)

Traditionally the introduction to proof in school has been via Euclidean geometry. However, this disappeared from the syllabus in Britain with the arrival of the new mathematics, and the NCTM standards suggest a change in emphasis in the USA, with increased attention recommended for:

- the development of short sequences of theorems
- deductive arguments expressed orally and in sentence or paragraph form

and decreased attention to

- Euclidean geometry as a complete axiomatic system
- Two column proofs.

The reasons for this are not hard to find. Senk (1985) showed that only thirty percent of students in full-year geometry courses reach a seventy five percent mastery on a selection of six geometric proof problems.

Not only is Euclidean proof hard, it fails to satisfy stringent tests of modern mathematical rigour because it depends on subtle intuitive notions of space. As Hilbert put it most succinctly:

> One must be able to say at all times - instead of points, straight lines and planes - tables, chairs and beer mugs.

But proof in terms of tables, chairs and beer mugs requires a great deal of sophistication which is not available to younger students. Mathematical proof as a human activity requires not only an understanding of the concept definitions and the logical processes, but also insight into how and why it works. Tall (1979) asked first year university students to comment on their preferences for the standard proof that $\sqrt{2}$ is irrational by contradiction, or for an alternative proof that showed the square of a whole number always had an even number of prime factors, hence the square of any fraction could not be $\frac{2}{1}$, because the prime 2 appeared an odd number of times. The students significantly preferred the second proof, because it gave some kind of *explanation* as to why the result was true (even though it was expressed in slightly loose mathematical terminology). In the transition to advanced mathematical thinking, mathematical insight in proof may be more important than mathematical precision.

Yet it does not take long before creases in the mind begin to form. Vinner (1988) gave students two proofs of the mean value theorem (if a function f is differentiable between *a* and *b* and continuous at *a* and *b*, then there is a point ξ between *a* and *b* such that $f'(\xi)=(f(b)-f(a))/(b-a)$.

(1) the standard algebraic proof (applying Rolle's theorem to $f(x) - \frac{f(b)-f(a)}{b-a}(x-a)$)

(2) a visual proof moving the secant *AB* in figure 5 parallel to itself until it becomes a tangent.

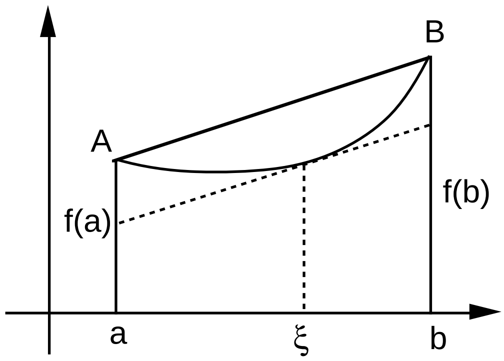


Figure 5 : A visual "proof" of the Mean Value Theorem

29 students found the visual proof more convincing, 28 the algebraic proof, and 17 considered them equal. It would be pleasing to see some evidence that the students criticized the geometric proof on some valid ground, for instance that it fails to give a proper algorithm to find ξ to any degree of accuracy. But almost all of those preferring the visual proof mentioned that it is "clear", "evident", "simple", etc whilst those preferring the algebraic proof tended to make general remarks that something is wrong or illegal with the visual approach. Vinner considers that students develop an "algebraic bias" not because of improved understanding of the algebra, but because of "habits, routines, convenience and metacognitive ideas which are 'environmental', not 'cognitive'". In particular he cites the current teaching of Euclidean geometry for sowing the seeds that a visual proof is unsatisfactory.

In mathematical analysis the need for a formal proof so often seems to arise out of fear that something might go wrong rather than confidence that something is right. To have a good intuition of what is right, one needs appropriate experiences to give a complete range of possible concept images and these are generally absent in undergraduates. On the other hand, students with experiences of magnifying a graph might realize that a graph may have tiny wrinkles on it – a positive reason why the smooth picture in figure 5 may not tell the full story – thus leading to the need for the formal proof.

Proof is concerned:

> ... not simply with the formal presentation of arguments, but with the student's own activity of arriving at conviction, of making verification, and of communicating convictions about results to others. Bell (1978)

In the last decade or so there has been a growing change of emphasis from teaching the *form* of proof, to encouraging the *process*, including the earlier stages of assembling information, specializing, generalizing and making and testing conjectures. Mason *et al* (1982) have developed a problem-solving approach in which the student builds up confidence by growing levels of conviction in a conjecture they have formulated:

convince yourself
convince a friend
convince an enemy.

The first of these requires the student to state a conjecture in a way which seems to him or her to be true, the second requires it to be articulated in a way which it can be meaningfully conveyed to someone else, and the third requires the argument to be clarified and organized in a way which will satisfy the meanest of critics. Nevertheless, this sequence of events stops short at what most professional mathematicians mean by proof: the logical deduction of theorems from carefully formulated concept definitions.

Alibert (1987, 1988a, 1988b) and his colleagues at Grenoble University have developed a course on Analysis, which include this final step through "scientific debate" in the classroom. This is based on the idea that students construct their own knowledge through "interactions, conflicts and re-equilibrations" and that the need for proof is best emphasized through making the contradictions explicit and involving the students in their resolution (Balacheff 1982). Rather than simply present a sequence of lectures in a logical sequence, followed by stereotyped exercises, the students (in a class of about 100) are encouraged to make conjectures. For instance, after the introduction of the notion of integral, the teacher might say:

> If I is an interval on the reals, a a fixed element of I, and x an element of I, we set, for f integrable over I,
>
> $$F(x)= \int_a^x f(t)\ dt\ .$$
>
> Can you make some conjectures of the form:
>
> if f ... then F ...

About twenty conjectures were formulated by the students, such as:

> "if f is increasing then F is too."

(which happens to be false) are then considered for debate. Arguments in support or against these conjectures are then addressed to the other students in a way which must convince everyone (including the speaker).

75% of students responding to a questionnaire preferred the method of incorporating debates whilst 10% rejected it as being inaccessible and not sufficiently organized. Many find debates helpful particularly when new ideas are introduced, but prefer the teacher to round off the debate by summarizing the knowledge gained.

Thus successful methods are being developed in mathematics education to improve the students participation in the processes of mathematical thinking, including the necessity for precise definitions and logical deduction. But these methods depend on radically different approaches on the part of teachers and only time will tell if they will become more widely accepted.

Reflections

Looking back over the evidence assembled, there is a great deal of data to support the existence of serious cognitive conflict in the learning of more advanced mathematical processes and concepts such as functions, limits, infinity and proof. What also seems to be clear is that the formal definitions of mathematics, that are such effective foundations for the logic of the subject, are less appropriate as cognitive roots for curriculum development. Their subtlety and generality are too great for the growing mind to accommodate all at once without a high risk of conflict caused by inadvertent regularities in the particular experiences encountered. There are creases of the mind everywhere: in teachers, in professional mathematicians, in mathematics educators, as well as in students. Given such a catalogue of difficulties, is there a way ahead?

We should not be too downhearted. The mathematical culture of which we speak is the product of three thousand, or is it three million, years of corporate human thought. It is asking a great deal to compress such diverse richness of experience into a decade or so of an individual's schooling. What is certain is that if we try to teach these ideas without taking account of the cognitive development of the student then we will surely fail with all but the most intelligent - and even these will have subtle creases in their minds as a result of their experiences. It is essential therefore that the expert be willing to re-examine his or her beliefs in the nature of mathematical concepts and be prepared to see them also from the viewpoint of the learner.

So much of the research quoted in this chapter has been built on implicit unspoken assumptions about the nature of the concepts being considered. The first step on a research agenda to assist students in the transition to advanced mathematical thinking must therefore be the clarification of these unspoken assumptions and the sensitizing of researchers and teachers to their existence. One source of evidence for them is in clinical interviews with students and a careful reflection on what is said, not just to see how it conflicts with formal mathematics, but also to place formal mathematics itself into perspective as a human activity which attempts to organize the complexities of human thought into a logical system. A theory of cognitive development of mathematical thought in the individual, from elementary beginnings through to formal abstractions, requires a cognitive understanding of the formal abstractions themselves.

The second stage is for more detailed clinical observations of the transition process as the massive process of cognitive restructuring takes place. This transition involves a number of difficult cognitive changes:

- from the concept considered as a process (the function as a process, tending to a limit, potential infinity),

- to the concept encapsulated as a single object that is given a name (the function as an object, the limit concept, actual infinity),

- via the abstraction of properties to the concept given in terms of a definition (function as a set of ordered pairs, the epsilon-delta limit),

- to the construction of the properties of the defined object through logical deduction,

- and the relationships between various different representations of the concept (including verbal, procedural, symbolic, numeric, graphic).

These are not intended to represent a hierarchy of development, in particular the relationships between various representations permeate the whole system in a horizontal manner whilst the conceptual strands become more sophisticated. Empirical evidence traditionally suggests that it is necessary to become familiar with a process before encapsulating it as an object. The computer is capable of carrying out routine processes (such as drawing graphs) which now give the possibility of new learning strategies in which the objects produced by the computer are the focus of attention before the internal algorithms are studied.

The third stage is then the design and testing of learning sequences aimed at assisting the cognitive reconstruction in the transition to advanced mathematical thinking.

Research to date aimed at improving learning – as opposed to research which simply observes what is currently occurring – has a common thread. This is that true progress in making the transition to more advanced mathematical thinking can be achieved by helping students to reflect on their own thinking processes and to confront the conflicts that arise in moving to a richer context where old implicit beliefs no longer hold. Such intellectual growth is stimulated by flexible environments which furnish appropriate cognitive roots and help the student to build a broader concept image. Over-simplified environments designed to protect students from confusion may only serve to provide implicit regularities that students abstract, causing serious conflict at a later stage.

In taking students through the transition to advanced mathematical thinking we should realize that the formalizing and systematizing of the mathematics is the final stage of mathematical thinking, not the total activity. As Skemp wrote in the "Psychology of Learning Mathematics" (1971):

> Some reformers try to present mathematics as a logical development. This approach is laudable in that it aims to show that mathematics is sensible and not arbitrary, but it is mistaken in two ways. First it confuses the logical and the psychological approaches. The main purpose of a logical approach is to convince doubters; that of a psychological one is to bring about understanding. Second, it gives only the end-product of mathematical discovery ('this is it, all you have to do is learn it'), and fails to bring about in the learner those processes by which mathematical discoveries are made. It teaches mathematical thought, not mathematical thinking.

In like manner, at the advanced level, teaching definitions and theorems only in a logical development teaches the product of advanced mathematical thought, not the process of advanced mathematical thinking.

References

Alibert D. (1987). Alteration of didactic contract in codidactic situation. *Proceedings of the 11th International Conference of PME*, Montréal, 379-385.

Alibert D. (1988a). Towards new customs in the classroom. *For The Learning of Mathematics*, 8, 2, 31-43.

Alibert D. (19880. Co-didactic system in the course of mathematics: how to introduce it.*P.M.E. 12, Hungary*, 109-116.

Artigue M. (1986). The notion of differential for undergraduate students in the sciences. *P.M.E.10, London*, 235-240.

Ayers T., Davis G., Dubinsky E. & Lewin P. (1988). Computer experiences in learning composition of functions. *Journal for Research in Mathematical Education*, 19,3, 248-59.

Balacheff N. (1982). Preuve et démonstration. *Reserche en Didactique des Mathématiques*, vol 3. 3.

Barnes M (1988). Understanding the Function Concept: Some Results of Interviews with Secondary and Tertiary Students. *Research on Mathematics Education in Australia*, 24-33.

Bell A.W. (1978). The learning of process aspects of mathematics. *Proceedings of the second International Conference for the Psychology of Mathematics Education*, Osnabrück, 48-78.

Borasi, R. (1985). Intuition and rigor in the evaluation of infinite expressions. *Focus on Learning Problems in Mathematics*, 7, 3-4, 65-75.

Buck R. C. (1970). Functions. In E.G. Begle (Ed.). *The sixty-ninth yearbook of the National Society for the Study of Education, Part I: Mathematics Education*. Chicago, IL: National Society for the Study of Education.

Breidenbach D., Dubinsky E., Hawks J, and Nichols D. (*in preparation*). Development of the Process Concept of Function.

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.

Cornu B. (1983). *Apprentissage de la notion de limite: Conceptions et Obstacles*. Thèse de Doctorat, Grenoble.

Davis R.B. (1982). *Learning Mathematics: The Cognitive Science Approach to Mathematics Education*. Croom-Helm, London.

Davis R.B. and Vinner S. (1986). The Notion of Limit: Some Seemingly Unavoidable Misconception Stages. *Journal of Mathematical Behaviour*, 5, 3, 281-303.

Demana & Waits 1988. Pitfalls in graphical computation, or why a single graph isn't enough. *College Mathematics Journal*,19 (2) 177-183.

Desmond A.J. (1975). *The Hot-Blooded Dinosaurs*. Blond & Briggs, London.

Dorofeev, G. V. (1978). The concept of function in mathematics and in school. *Mathematics in School*, 2, 10 -27.

Dreyfus T. and Eisenberg T. (1982). Intuitive functional concepts: a Baseline Study on Intuitions. *Journal for Research in Mathematical Education*, 6,2, 18-24.

Dreyfus T. and Eisenberg T. (1987). On visualizing function transformations, (to appear)
Dreyfus T. and Vinner S. (1982). Some aspects of the function concept in college students and junior high school teachers. *Proceedings of the Sixth International Conference for the Psychology of Mathematics Education*, Antwerp, 12-17.
Dreyfus T. & Vinner S. (1989). Images and Definitions for the Concept of Function. *Journal for Research in Mathematics Education*, *20*(4), 356-366.
Dubinsky E. (1990). Computers in teaching and learning discrete mathematics and abstract algebra. *Advanced Technologies in the Teaching of Mathematics and Science* (ed. D. L. Ferguson). Springer-Verlag. (to appear).
Dugdale (1982). Green globs: A microcomputer application for graphing of equations. *Mathematics Teacher*, *75*, 208-214.
Dugdale S. & Kibbey D. (1989) Building a qualitative perspective before formalizing procedures: graphical representations as a foundation of trigonometric identities, *Proceedings of the 11th Meeting of PME-NA*. NJ.
Ervynck G. (1981). Conceptual difficulties for first year students in the acquisition of the notion of limit of a function, *Actes du Cinquième Colloque du Groupe Internationale PME*, Grenoble, 330-333.
Even R. (1988). *Prospective secondary mathematics teachers' knowledge and understanding about mathematical function*, unpublished Ph.D. thesis, Michigan State University.
Fischbein E. (1978). Intuition and mathematical education, *Osnabrücker Schriften zür Mathematik 1*, 148-176.
Fischbein E., Tirosh D., Hess P. (1979). The intuition of infinity, *Educational Studies in Mathematics* 12, 491-512.
Fischbein E., Tirosh D., Melamed U. (1981). Is it possible to measure the intuitive acceptance of a mathematical statement?, *Educational Studies in Mathematics*, 12 491-512.
Goldenberg E. P. , Harvey W., Lewis P.G., Umiker R.J., West J. & Zodhiates P. (1988). *Mathematical, technical and pedagogical challenges in the graphical representation of functions* (Tech. Rep. No.88-4), Educational Technology Center, Harvard Graduate School of Education.
Janvier C. (1987). Representations and understanding: The notion of function as an example. In C. Janvier (ed.), *Problems of representation in mathematics learning and problem-solving.* Hilsdale, NJ: Lawrence Erlbaum Associates.
Keisler H.J. (1976). *Elementary Calculus*, Prindle, Weber & Schmidt, Boston.
Kerslake D. (1977). The understanding of graphs, *Mathematics in School*, 6,2 22-25.
Kleiner I. (1989). Evolution of the function concept: A brief survey. *The College Mathematics Journal*, 20 (4), 282–300.
Leinhardt G, Zaslavsky O, Stein M. (1990). Functions, Graphs, and Graphing: Tasks, Learning and Teaching. *Review of Educational Research*, *60* 1 1-64.
Malik M. A. (1980). Historical and pedagogical aspects of the definition of a function. *International Journal of Mathematics Education in Science and Technology*, 11(4), 489-492.
Mason J., Burton L. & Stacey K. (1982). *Thinking Mathematically*, Addison Wesley.
Markovits Z., Eylon B. & Bruckheimer M. (1986). Functions today and yesterday, *For the Learning of Mathematics*, 6, 18-24.
Markovits Z., Eylon B. & Bruckheimer M. (1988). Difficulties Students have with the Function Concept, *The Ideas of Algebra, K-12*, N.C.T.M. 1988 Yearbook, 43-60.
Monaghan J. D. 1986. *Adolescent's Understanding of Limits and Infinity*, unpublished Ph.D. thesis, Warwick University.
Orton A. (1977). Chords, secants, tangents & elementary calculus, *Mathematics Teaching*, 78 48-49.
Orton A. (1980). *A cross-sectional study of the understanding of elementary calculus in adolescents and young adults*, unpublished Ph.D., Leeds University.
Papert S. (1980). *Mindstorms*, Harvester Press, Brighton.
Rich A., Rich J & Stoutemyer D.(1989). *Derive™ - A Mathematical Assistant*, Soft Warehouse, Honolulu.
Robert A. (1982). L'Acquisition de la notion de convergence des suites numériques dans l'Enseignement Supérieur, *Reserches en Didactique des Mathématiques*, vol 3, no 3, 307-341.
Schaaf W.A. (1930). Mathematics and World History. *Mathematics Teacher* 23 496-503.
Schwarzenberger R.L.E. & Tall D.O. (1978). Conflicts in the learning of real numbers and limits, *Mathematics Teaching* 82, 44-49.
Schwartz J. L. (1990). Software to Think With: The Case of Algebra, *Advanced Technologies in the Teaching of Mathematics and Science* (ed. D. L. Ferguson). Springer-Verlag. (to appear).
Schwartz J.T., Dewar R.B.K., Dubinsky E. & Schonberg E. (1986). *Programming with Sets, An introduction to SETL*, Springer-Verlag.
Schwarz B. & Bruckheimer M. (1988). Representations of functions and analogies, *P.M.E. XII, Hungary*, 552-559.
Schwarz B., Dreyfus T. & Bruckheimer M. (1988). *The Triple Representation Model Curriculum for the Function Concept* (to appear).

Senk S. L. (1985). How well do students write geometry proofs?, *Mathematics Teacher*, 78,6, 448-456.

Sierpiń ska A. (1985a). Obstacles epistémologiques relatifs à la notion de limite, *Reserches en Didactique des Mathématiques*, 6.1, 5-67.

Sierpiń ska A. (1985b). La notion d'un obstacle epistémologique dans l'enseignement, *Actes de la 37e Rencontre CIEAEM*, Leiden.

Sierpiń ska A. (1987). Humanities students and Epistemological Obstacles Related to Limits, *Educational Studies in Mathematics*, 18,4, 371-87.

Sierpiń ska A. (1988). Epistemological remarks on functions, *P.M.E. XII, Hungary*, 568-575.

Skemp R.R. (1971). *The Psychology of Learning Mathematics*, Pelican, London.

Sullivan K. (1976). The teaching of elementary calculus: an approach using infinitesimals, *American Mathematical Monthly* 83,5 370-375.

Taback S. (1975). The child's concept of limit in *Children's Mathematical Concepts* (ed. Roskopff M.F.) Teachers' College Press, New York 111-114.

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

Tall D.O. (1979). Cognitive aspects of proof, with special reference to the irrationality of √2, *Proceedings of the Third International Conference for the Psychology of Mathematics Education*, Warwick, 203-205.

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

Tall D.O. (1980b). Mathematical intuition, with special reference to limiting processes, *Proceedings of the Fourth International Conference for the Psychology of Mathematics Education*, Berkeley 170-176.

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

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

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

Tall D.O, Blokland P. & Kok D. (1990). *A Graphic Approach to the Calculus*, (for I.B.M. compatible computers), Sunburst.

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

Thomas H.L. (1975). The concept of function. In M. Rosskopf (ed.) *Children's mathematical concepts*. New York: Teachers College, Columbia University.

Tirosh D. (1985). *The intuition of infinity and its relevance for mathematics education*, unpublished doctoral dissertation, Tel-Aviv University.

Vinner S. and Hershkowitz R. (1980). Concept images and common cognitive paths in the development of some simple geometrical concepts, *Proceedings of the Fourth International Conference for the Psychology of Mathematics Education*, Berkeley, California, 177-184.

Vinner S. (1982). Conflicts between definitions and intuitions - the case of the tangent, *Proceedings of the 6th International Conference of P.M.E., Antwerp*, 24-28.

Vinner S. (1983). Concept definition, concept image and the notion of function, *The International Journal of Mathematical Education in Science and Technology*, 14, 293-305.

Vinner S. (1988). Visual Considerations in College Calculus - Students and Teachers, *Theory of Mathematics Education, Proceedings of the Third International Conference, Antwerp*, 109-116.

Wheeler M.M. & Martin W.G. (1987). Infinity Concepts among pre-service elementary school teachers, *Proceedings of the Eleventh International Conference of P.M.E.*, Montréal.

Wheeler M.M. & Martin W.G. (1988). Explicit Knowledge of Infinity, *P.M.E.-N.A., Proceedings of the Tenth Annual Meeting*, Dekalb, Illinois.

Yerushalmy M. (1990). Understanding concepts in algebra using linked representation tools. *Advanced Technologies in the Teaching of Mathematics and Science* (ed. D. L. Ferguson). Springer-Verlag. (to appear).

Yerushalmy M. (1991) Students' Perceptions of Aspects of Algebraic Function Using Multiple Representation Software, *Journal of Computer Assisted Learning.* Blackwell Scientific Publications (to appear).