Differing Modes of Proof and Belief in Mathematics

David Tall

Mathematics Education Research Centre
University of Warwick, UK
e-mail: david.tall@warwick.ac.uk

This presentation will consider different modes of proof that may be appropriate in mathematics and the kinds of belief that underpin them. In particular, it will distinguish between three distinct worlds of mathematical thought: the ***embodied****, the* ***proceptual*** *(using symbols as* ***pro****cess and con****cept*** *in arithmetic and algebra) and the* ***formal****. Embodied proof builds from our interaction with the outside world through our senses. It is capable of sophisticated development through language and human interaction to encompass subtle arguments about models of the world, such as Euclidean Geometry. Proceptual proof refers to the use of symbols in arithmetic and algebra (and the wider theory of procepts in which symbols operate dually as process and concept) where a statement can be proved through calculation and manipulation. Formal proof is built by formal deduction from axioms and concept definitions to construct coherent mathematical systems. The presentation will make the case that each of these forms of proof inhabits its own mathematical world of thought with distinct criteria for belief and truth. This will be used to shed light on the mental constructions required for different modes of proof and on the range of cognitive tasks experienced by individuals in their quest for understanding.*

Introduction

This presentation gives me the opportunity to bring together a theory of the conceptual development of mathematical concepts that has been developing for several years (e.g. Tall, 1995, see Appendix). In particular I shall focus on what it has to tell us about the *process* of proving and the concept of *proof*.

Mathematical proof is a concept that arises in the work of mathematicians, and as such it has a sophisticated meaning. I can illustrate this with an example of how a mathematician might think of a simple everyday concept such as 'two'. In a long conversation I had with the Greek mathematician Negrepontis at a mathematics education conference, he alighted on the meaning of the word 'two' according to his hero Plato. He explained that Plato was unsatisfied with the worldly meaning such as 'two stones' or 'two oxen' because the stones or oxen were different things and twoness essentially involved having two identical things. He therefore began a long discourse on how one could invent two things which were the same, yet had features which allowed one to distinguish between them. I caught the sense of his argument from a philosophical point of view but I tried to put forward that the idea of 'twoness' in the growth of a child could not be approached from such a lofty intellectual position. He continued at length and, at a moment when he drew breath, I interjected and said that, if he would let me make one comment in one sentence that I felt was relevant, then I would remain silent so that he could complete the burden of his argument. After some protestation from him, I got in my one sentence. It was this: 'Plato was very young when he was born.'

Negrepontis continued on his discourse about the 'golden section', which he explained could be computed by what we would nowadays recognize as an infinite series in two essentially different ways. These different sequences of computation had the same eventual value, and therefore we had two things that were the same (in value) but essentially different. With a great final flourish he ended up with Plato's idea of twoness involving two objects that were the same, but could be distinguished.

At this point I reminded him of my one sentence interjection. I acknowledged that Plato was indeed highly sophisticated in his meaning for the number two, but that, when he was born, Plato could not talk and lived for months off his mother's milk without a philosophical discussion of any kind. My question to him was then, 'How did Plato, who was once a helpless child, develop into the great philosopher whose name has lasted for over two millennia?' Surely the answer lies in some kind of cognitive development, and what education is interested in is not just the product of that development, but the *process* of how it occurs and can be encouraged to occur.

My discussion today grows out of this observation. I ask: 'How does the sophisticated notion of mathematical proof arise in the development of the child and how can we encourage it to occur?' To begin to respond to this question I need to begin with the notion of proof itself.

What is proof for a mathematician?

Proof is one of those fundamental concepts on which all mathematicians seem to agree. Yet if we dig a little beneath the surface, we find subtleties that are not easily explained away. In their wonderful book, *The Mathematical Experience*, Davis and Hersh caricature what they call 'the ideal mathematician':

> He rests his faith on rigorous proof; he believes that the difference between a correct proof and an incorrect one is an unmistakable and decisive difference. He can think of no condemnation more damning than to say of a student: 'He doesn't even know what a proof is.' Yet he is able to give no coherent explanation of what is meant by rigor, or what is required to make a proof rigorous. (Davis & Hersh, 1981, p. 34)

The authors continue with fictional interchanges between the 'ideal mathematician' and representatives typical of other groups. The ideal mathematician is unable to give any explanation to a university's public information officer of the details or applications of his work, or say anything that is understandable to 'the ordinary citizen'. (p. 38.) He fares no better with a student, explaining that 'proof is what you've been watching me do at the board three times a week for three years'. (p. 39.) … 'everybody knows what proof is. Just read some books, take courses […], you'll catch on.' (p. 40.) He has difficulties with a philosopher too, ending a conversation by saying 'I'm not a philosopher, philosophy bores me. You argue, argue, argue and never get anywhere. My job is to prove theorems, not to worry about what they mean'. (p.41).

In this way we see that, though proof is the central idea of modern mathematics, quite what it *is* is a matter of implicit, rather than explicit, agreement between members of the mathematics community. In my own case, as a first year undergraduate, I was intoxicated with the beauty of Halmos's *Naïve Set Theory* and looked forward to the day when the naïve viewpoint expressed in this book was supplanted by the ultimate formal truth. For me, that truth never came. I believe now that this perfect proof is an illusion. It is sensed by communities[1] of mathematicians who strive for such perfection, but have to be satisfied with a compromise that stylises proof to focus on important issues while suppressing those things that are considered contextual.

I realised this myself intuitively when I wrote *Foundations of Mathematics* with Ian Stewart over 25 years ago (a book still in use in undergraduate mathematics today). There we spoke not only of the need for explicit hypotheses in quantified set-theoretic statements, and the deduction of successive lines of proof using only axioms, established theorems in the theory, and preceding lines of the current proof, but also of the 'contextual truths' that had been long established, so that they need no longer be quoted explicitly to burden the reader. As an example, we built up over several chapters, the notion of the real numbers as a complete ordered field in three stages:

1. proving sufficient properties directly from the axioms for a field to establish a context for arithmetic.
2. Assuming properties of arithmetic as being contextual and focusing on properties of order,
3. Assuming properties of arithmetic and order as being contextual and focusing on the property of completeness.

In this way we attempted to introduce young undergraduates to the style of presentation of mathematics, distinguishing between what was, in a given context, self-evident and not in need of specific comment and what was new and needed explicit proof. These ideas of what may be assumed contextually

[1] Here I say 'communities' in the plural because, although mathematicians would claim to generally agree on standards for proof , different communities, such as topologists and logicians or constructive mathematicians and non-standard analysts, may present proofs in distinctly different ways, including unspoken assumptions that they take for granted.

and what needs to be explicitly proved, using logical deduction and previously established results, is highly non-trivial and, I would suggest, is implicit rather than explicit in the minds of most mathematicians.

This is a natural consequence of how the brain focuses on essential ideas and suppresses less significant material. It happens as we learn new ideas in mathematics and lose contact with earlier ideas that are no longer sufficiently relevant or powerful to take up thinking space. In *The Learning of Mathematics 8-13*, we find the comment:

> After the formalization had been taught, or three months later, the practical or pre-formalization work which led up to it was often forgotten or not seen as significant.
> (Johnson (ed.), 1989, p. 219)

In his article on *Proof and Progress in Mathematics*, Fields medallist Thurston observed:

> After mastering mathematical concepts, even after great effort, it becomes very hard to put oneself back into the frame of mind of someone to whom they are mysterious.
> (Thurston, 1994, p. 947)

Such forgetting has interesting implications for experts planning the curriculum to teach students:

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

In this way we may conjecture that most professional mathematicians have become so expert in the conventions of their community that they are no longer aware of how they came to be so sophisticated. At a conference focused on proof, it is therefore incumbent upon us to revisit the notion of proof, to seek its special characteristics, dependent in part on the community of mathematicians, but also to investigate the cognitive growth of proof as the individual matures to find out how we might hope to encourage our students to grasp its essential nature. That nature, however, cannot be found simply by attempting to reproduce what it is that mathematicians do in a form appropriate for younger and more naïve students. We need to look a little deeper into the nature of proof itself, not only within the community of practice of mathematicians, but in its human development.

In Tall (1999) I made a foray into the cognitive development of proof, considering how proof is dependent upon the representations that the student has available, starting with prediction and experiment in the real world, proof by computation in arithmetic and manipulation in algebra, proof at university. Subsequently, I realised the fundamentally different ways of thinking in different modes of mathematical representation, which relate not just to formal *proof*, but to human *belief*.

Three worlds of mathematics

Over recent years I have been attempting to grasp the essential nature of cognitive growth from child to mathematician, including how and why different individuals develop mathematical thinking in different ways. I realised slowly that the lines of development of my work were coalescing into three distinct threads:

1. Those relating to our sensory perceptions of and physical actions on the real world and our reflections on these perceptions and actions that led to conceptions of properties then relationships between properties that I began to realise led naturally to the individual to *invent* Platonism.
2. Those relating to our use of symbolism in arithmetic, algebra and more general analytic forms that enabled us to calculate and manipulate to get answers.
3. The formal axiomatic world of mathematicians that is the final bastion of presentation of coherent theories and logical proof using axioms and definitions expressed in quantified set-theoretical statements that could be manipulated using the laws of logic.

Over the last decade and more, these threads have crystallised into a form where I see them as quite different worlds of mathematics, with different ways of operating, developing different standards of validity and truth. I have named them, and characterised them as follows:

1. The *embodied world* of perception and action, including reflection on perception and action, which develops into a more sophisticated Platonic framework,
2. The *proceptual world* of symbols, such as those in arithmetic, algebra and calculus that act as both *pro*cesses to do (e.g. 4+3 as a process of addition) and concepts to think above (e.g. 4+3 as the concept of sum) as formulated in the theory of procepts (Gray and Tall, 1994).
3. The *formal world* of definitions and proof leading to the construction of axiomatic theories, (Tall, 1991).

For a more detailed discussion of these ideas, consult Tall 2002, Watson, et al, 2002, available as downloads from my website www.davidtall.com/papers.

I am aware of such previous uses of the term 'world' such as microworlds, or the 'three worlds of Popper'. At first I referred to them in other ways, perhaps as 'modes of thought' or 'stages of development'. However, as I looked at a wide range of contexts in which we think about mathematics, the constant distinctions between them made me realise even more that, whilst some individuals can operate and live in each of these modes as they see fit, many operate predominantly in one world (or perhaps two). I saw how different the worlds of proof in geometry and proof in arithmetic and algebra are, and the huge difficulty most university students have with the transition to the formal world. So, for me, the notion of 'three different worlds' of mathematics has become a most useful and productive way of thinking.

Cognitive development of the three worlds

The three worlds develop cognitively in sequence, first the embodied world in a sensori-motor form, with its roots in physical perception and action. Then the proceptual world begins by building from embodied actions such as counting, adding, grouping and sharing, to develop symbolic forms for number, sum, product, division, and so on. This world is characterised by the use of symbols that operate dually as process (e.g. counting) and concept (e.g. sum). As the child grows, both these worlds are available. It is much later, if at all, that individuals meet the formal world of axiomatic definition and proof.

Each world develops in an idiosyncratic way. The embodied world is initially based on perception of the properties of physical objects. Language allows us to furnish them with descriptions, which allows us to identify them – that is a circle 'because it is round', that is a square' because it has four equal sides and its angles are right-angles'. Subtly the descriptions become more precise until the purpose is shifted to *definitions* of objects, where we can now test if an object is what we claim it is by checking if it satisfies the definition. As soon as an object has a definition, say an isosceles triangle is 'a triangle with two sides equal' then we can see if this implies that it has other properties, for instance that it is 'a triangle with two *angles* equal'. Definitions of concepts naturally lead to a process of deduction: *if* the object has these properties, *then* it will have those properties. In this way, using the idea of 'congruent triangles' to establish a template for 'sameness', Euclidean proof is developed as a verbal formulation of embodied properties of geometric figures. This leads to the concept of Euclidean theorem, and such theorems can be built, one upon another to develop the theory of Euclidean geometry (van Hiele, 1986).

The proceptual world of symbols is quite different. Each concept starts out as a process, for example the process of counting uses symbols that allow us also to think of them as (number) concepts. Likewise, the process of addition 3+4 becomes the concept of sum, the process of repeated addition (multiplication) becomes the concept of product. This leads to the calculations of arithmetic, carried out initially as (embodied) processes becoming symbolic algorithms. From here, algebra develops, in part as an expression of the generalization of arithmetic, where a symbol such as $2n-1$ represents a typical process of evaluation, in this case "double n and take away one". This process (of evaluation) is encapsulated as a concept (of expression), the expression $2n-1$, which can now be manipulated by algebraic operations.

The formal world uses such experiences from both these worlds to build the world of formal definition and proof. However, there is an important difference between Euclidean proof and formal proof. In the embodied world we 'know' objects through our senses and select properties to define them. In the formal world we first make the definitions using quantified set-theoretic statements and then construct objects that satisfy them. This move from the embodied 'concepts giving rise to definitions' to the formal 'definitions giving rise to concepts' is a change in emphasis that involves a significant reconstruction of meaning. In addition there is a change in practice from the use of (embodied) 'prototypical figures' in Euclidean geometry to quantified statements that causes a seismic shift in meaning. Formal proof as a *process* begins from explicit quantified definitions and deduces that other properties hold as a consequence. This establishes theorems as *concepts* in their own right, which can be used as building blocks to build yet more theorems, constructing a logically deduced formal theory from explicit foundations.

Validity in different worlds of mathematics

Proof, which is the central idea of the formal world, has earlier manifestations in the embodied and proceptual world, each of which has a distinct notion of validity. However, mathematicians rarely endow these earlier forms of argument as being worthy of the name 'proof'. An alternative approach is to consider not only the process of deduction, but also wider notions of validity that may be used by students to justify their beliefs. One such notion is a 'warrant for truth'.

I first met this idea framed in the notion of 'mathematical warrant' in the PhD thesis of Melissa Rodd (1998, 1997). Rodd (2000) defines a warrant as 'that which secures knowledge'. As one of several examples, she considers the case that the sum of the first n whole numbers is $\frac{1}{2}n(n+1)$ and discusses whether various experiences can act as a warrant. These include:

- proof by induction,

and

- a picture of the sum of the first n numbers as a staircase, and the putting together of two staircases to give a rectangle of area n by n+1 (figure 1).

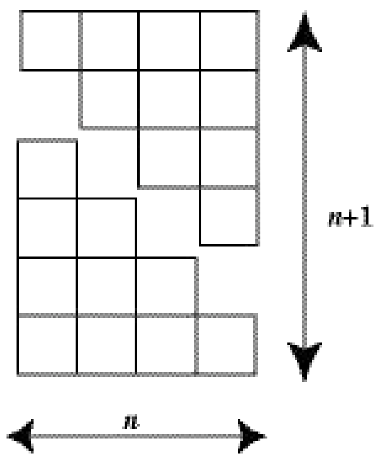

Figure 1: twice the sum 1+…+n is a rectangle n by n+1.

In addition to these two proofs, I would add the intermediate proof in the proceptual world, as essentially discovered by the boy Gauss in adding the numbers 1 to 100.

$$
\textit{add columns} \downarrow \quad
\begin{array}{rrrrr}
1 + & 2 + & \ldots + & 99 + & 100 \\
100 + & 99 + & \ldots + & 2 + & 1 \\
\hline
101 + & 101 + & \ldots + & 101 + & 101
\end{array}
= 100 \times 101
$$

Figure 2: twice the sum of the first 100 numbers is 100 × 101.

He simply imagined the numbers in the reverse direction and, on adding them in succession, he obtained 101 for each sum a total of 100 times, so twice the sum is 100×101 and the sum of the first 100 numbers is half of this. The proof is generic and applicable to any number of terms.

These examples offer three ways of convincing:

(a) the pattern in a picture (embodied),

(b) the pattern in an arithmetic calculation (proceptual),

(c) an induction proof (formal).

The status of these warrants for truth is dependent on the community of practice. For many students (and for Rodd herself), (a) and (b) are convincing, but (c) is not. For the community of mathematicians, however, (c) is a proof, and (a) and (b) may have deficiencies and may be relegated to the level of 'arguments'. For instance, (b) is not a proof if it only deals with a single value (in the given case, n = 100) but if it is seen as a generic case that applies to all natural numbers, then it is seen as being more convincing.

What this shows are three distinct ways of convincing, one embodied, one proceptual and one formal. Each has a different warrant, the embodied one uses a pattern to allow us to 'see' the truth of the statement, the second uses the same pattern in a numerical calculation, the third uses a formal proof by induction.

However, the situation is more complex than this. Each world of mathematics has its own sequence of development and warrants for truth are not static notions that are fixed for all time. As the individual grows in sophistication, the need for subtler arguments becomes apparent.

Different sequences of development for proof

The embodied world encourages a growth of subtlety in proof from initial experiments with the physical world to later proofs in Euclidean geometry. In the embodied world, warrants for truth begin through physical prediction and experiment to test whether the experiment leads to an expected conclusion. How do we 'know' that the angles of a triangle add up to 180°? Well, just cut a triangle out of paper, and tear off its corners. When you put the three corners together, they form a straight line!

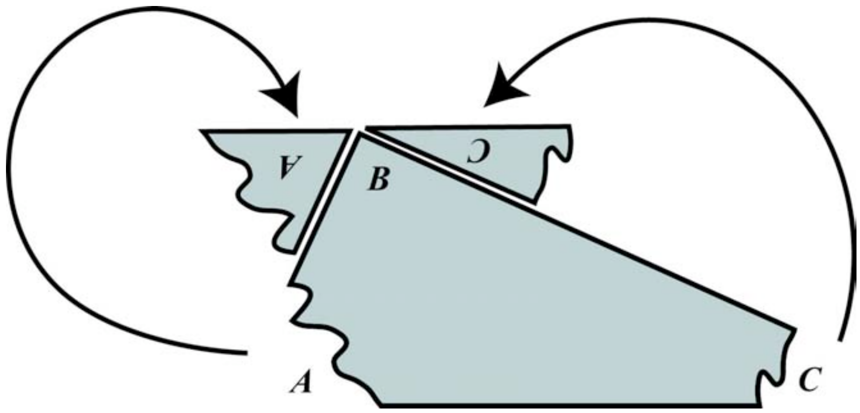

Figure 3: A physical demonstration

In this particular case, the demonstration involves tearing a specific triangle. However, this is a *generic* example that enshrines the general argument in the particular case. By performing a *thought experiment* the individual may later transform such an enactive, physical proof into a verbal Euclidean proof by reflecting on it. Suppose one took a triangle *ABC* and drew a line *DE* through *B* parallel to *AC*. Then the angles ∠DBA and ∠BAC are alternate angles between parallel lines and are therefore equal, similarly angles $\angle EBC$ and $\angle BCA$ are equal. But ∠DBA, ∠ABC and ∠CBA make a straight line, which is 180°, so ∠BAC, ∠ABC and ∠BCA also add up to 180°, as required.

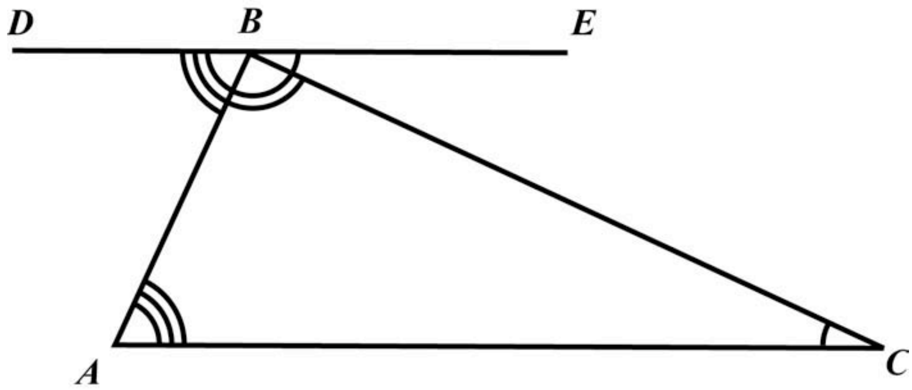

Figure 4: A Euclidean Proof

This example shows the embodied world as a natural environment to grow from physical experiment on figures through the use of language to build deductive relationships in Euclidean geometry. Indeed, the research of Rosch *et al*. (1976) on 'basic categories', shows that the child first names categories that are the most general that have an image to represent that category such as 'apple, doll, dog' rather than a more general category such as apple as a fruit or a more refined category of apple such as a Cox, a russet or a golden delicious. This is just what mathematicians do. They take basic categories of geometric figures as a starting point, such as triangle, circle, point, line, and build up a more comprehensive structure of formal relationships between them.

Generic examples using actions on specific objects can also give warrants for truth of arithmetic properties. For instance, we can *see* that two and three make the same total as three and two by holding up the requisite number of objects and the total is the same irrespective of the order. This 'order irrelevancy principle', later called 'the commutative law', arises through experience and observation. In the embodied world it is *not* a 'law' imposed *on* the real world, it is an *observation* of what happens *in* the real world.

We do not need to perform the calculation for *all* numbers, we can see the general statement in the particular pattern for 2+3 = 3+2, as in the (lengths of) arrays of blocks forming a number track:

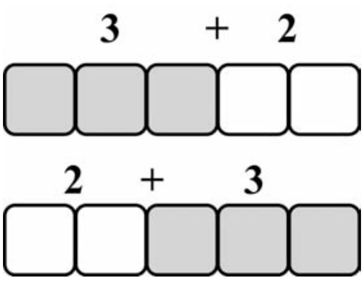

Figure 5: 3+2 is the same (length) as 2+3

It does not matter if we were to draw 3+2 or 5+7 or 12+9, they all follow the same kind of pattern. Again, a generic example gives a warrant for truth of the notion of order irrelevance of addition.

In the proceptual world, truth can also be tested by computation and manipulation. Our experiences of generic examples from the embodied beginnings of arithmetic give us warrants for truth of the general laws such as the commutativity of addition '$x+y=y+x$' or the distributivity of multiplication over addition '$x(y+z)=xy+xz$'. Children are just assumed to 'know' them to be true from experience of arithmetic and assume they will also naturally hold in algebra. There are other 'truths', however, such as

$$(a-b)(a+b)=a^2-b^2$$

that we 'show to be true' by carrying out simpler operations of algebra based on 'facts' already 'know' to be true. For instance:

$$(a-b)(a+b)=(a-b)a+(a-b)b=a^2-ba+ab-b^2=a^2-b^2.$$

We know from the work of Collis over 30 years ago that some children are happy that sums of small numbers are commutative and even agree that $x+y=y+x$ in algebra, but when confronted with a daunting sum involving large numbers, they are caught up with the need to *calculate*, and lose their confidence that such a large sum is commutative. Thus we *cannot* assume that children 'know' the basic rules in this way, nor that they can have the sophistication to know which rules are 'known' (whatever that means) and which 'need to be proved'.

In the nineteenth century, mathematicians were beguiled by such examples to such an extent that Peacock pronounced his law of 'algebraic permanence', asserting that laws that hold in one mathematical system would naturally hold in a larger system. He used this law to justify carrying the rules of arithmetic over to the rules of algebra. He was, however, in a word, *wrong*, The reverse is true, in any extension of a given system, there are *always* rules from the old system that no longer hold in the bigger system as the bigger system has more structure in it. For instance in the natural numbers, there is always a 'next' number, in rational numbers there is not; the real numbers are ordered, the complex numbers are not.

Tall et al (2001) show the many discontinuities in the expansion of number systems in secondary schools, from whole numbers to integers, from integers to rational numbers, the use of whole number powers, fractional powers, negative powers, infinite decimals, infinite limit processes and so on. How, in such a system is a child to maintain his or her bearings to be sure that what 'is true' remains true? Those who succeed in the proceptual world invariably become accustomed to carrying out such tasks supported by their embodied imagery and the guidance and approval of the teacher. They learn by convention what is contextual (and not requiring proof) such as the commutative law, and what is requires proof, such as the formula for the difference between two squares.

In the formal world, the situation changes again. Statements like '$x + y = y + x$' or '$x(y + z) = xy + xz$' are no longer true because of one's prior experience, they are true because *they are asserted to be true as axioms* for an axiomatic mathematical structure, say in a field or ring. The symbol $x + y$ need no longer carry with it any meaning of addition or involve any actual process of computation. What matters is that the structure obeys the given axioms related to the symbols and then, having proved a result solely by logical deduction from the axioms, one knows that this result is then true of *all* the examples in which the axioms are satisfied.

In figure 6 we represent some of these ideas diagrammatically. The picture is not meant to be final, or complete, for there is still much reflection to be performed. However, it shows the growth of the embodied world on the left, starting with perception and action, at a stage when generic examples act as warrants for truth through experiment. Early perceptions are clarified and refined by verbal discussion, leading to descriptions of already existing objects. These descriptions change to more precise definitions, so that descriptions become *pre*scriptions. If an object has the properties cited in the definition then this prescribes that the object is an example of the defined object. From here, different ways of describing objects naturally lead to a process of deduction: *if* the object has *this* property, *then* it has *that* property. A procedure of deductive proof is established based on the notion of congruent triangle. This leads from the *process* of Euclidean proof to the *concept* of a Euclidean theorem, and on to the construction of the full panoply of Euclid's books as an exemplar of a cogently deduced theory from specified starting assumptions.

In comparison to this, the development of the proceptual world is quite different. The process of counting is encapsulated as the concept of number and generic examples based on embodiment help the child to build a sense of the general properties of arithmetic. This leads to the use of letters as generalised arithmetic, which itself gives a new warrant for truth. For instance, to show that the sum of two consecutive odd numbers is a multiple of 4, take the first to be $2n-1$ where n is any whole number, then the next is $2n+1$ and their sum is

$$(2n-1)+(2n+1)=4n-1+1=4n$$

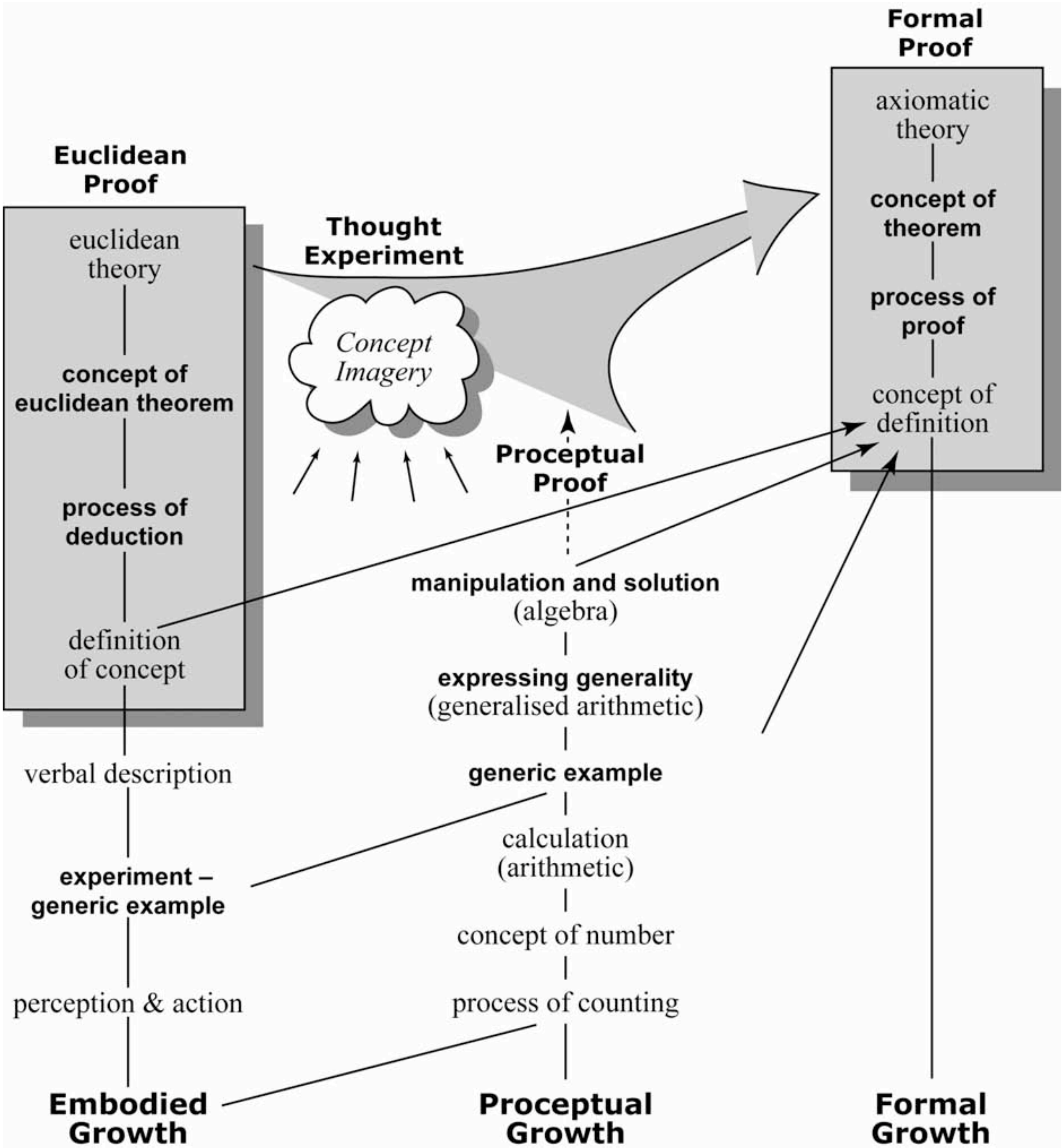

Figure 6: An outline of growth in three different mathematical worlds

which is clearly a multiple of 4. (But did you see which rules of arithmetic were used implicitly? Is it *really* proved from explicit truths? On the contrary, certain properties are used implicitly, such as the fact that a sum of several terms can be re-ordered and re-grouped without affecting the value.) Algebra gives us a warrant for truth, based on implicit use of ‘rules of arithmetic’, by manipulation of symbols. It also allows us to provide answers to problems by using algebraic methods of manipulation to produce a solution. But, in order to do this, one requires a sense of the conventions involved, which are usually implicit.

Moving to the formal world, we see the first item being the ‘concept of definition’, a neat play on words to distinguish the meaning from the phrase used in the embodied world, which was the ‘definition of concept’. This underlines the idea that the definition is used to give meaning to the concept in the formal world, not the other way round.

There is a parallel between the development of Euclidean proof in the embodied world and formal proof in the formal world. Indeed, in line with van Hiele theory, we see the definition leading to the process of proof, encapsulated as the concept of proof and then the concepts are organised in a deductive sequence to give the whole theory. This parallel was noted by Chin (2002) in his PhD thesis and will feature in another paper here (Chin & Tall, 2002). However, even here, when we speak of the sequence

definition → process of proof → concept of proof → axiomatic theory,

we find that the next category is developing before the previous one is established, so that a student can be at several points in the sequence at different times, as was found by Guttierez *et al* (1991) in van Hiele theory. This shows that, however we attempt to grasp the growth of proof and make it precise, the categories overlap and are not mutually exclusive.

Proof is not a purely formal art, whatever may be claimed by those who say they practice it. All new theorems must come from somewhere, and they do *not* come by writing down random statements to see what might or might not be proved. Mathematicians prove new theorems, not because they find them by formal proof, but because they have intuitions that intimate that certain theorems might be true, and *then* they set out to prove them. Take the case of *Fermat’s Last Theorem*, which took over three centuries to produce a proof of a ‘theorem’ that so many tried to prove and failed (see, for instance, Stewart & Tall, 2002).

Or, consider the case of John Nash, featured in the book and film *A Beautiful Mind*. There we find that the mathematicians at Princeton between the two great wars *despised* people who produced proof. Lefschetz, who was the head of department at the time, created an atmosphere in which what mattered was the ability to predict an intriguing theorem. If no one could prove it, so much the better. Such ideas are the stuff of mathematics. Once proved they are no longer interesting. Lefschetz wrote a famous book on Topology, which it was my pleasure to read as a graduate student at Oxford. However, he wrote it when he was on sabbatical, and this meant he had no students around to correct the many mistakes he made throughout the text! John Nash, too, valued 'good theorems' that others agreed were important but had not yet been proved. He used to set such theorems as exercises for his undergraduates on their examinations, asserting that they might, in their fresh innocence, see a path to the truth unavailable to those whose minds had been tarnished by too much effort in tackling them.

Thus it is that proof, the pinnacle of formal mathematics, is not achieved by formal means. So, how is it achieved? In figure 6, I add a cloud to represent the *concept imagery* that is part of the mental structure constructed by the individual in thinking about mathematics. Embodied ideas, sense of pattern in both embodiment and symbol use, all contribute to the concept image, (described in Tall & Vinner (1981), p.152 as 'the total cognitive structure that is associated with the concept, which includes all the mental pictures and associated properties and processes'). The concept image is used to imagine thought experiments, to conceive of possible definitions and possible theorems that might arise from those definitions. This gives a possible route from intuition to the formulation of theorems that might subsequently be given a formal proof.

Pinto and Tall (2001, 2002) show that such routes are available to undergraduates. Some successful undergraduates do take a *formal* route to learning about proof, accepting the definitions, committing them to memory and working with the logical relationships to build up a knowledge of formal mathematics. Others, however, take a *natural* route, building from their concept images and reconstructing their concept images in increasingly sophisticated ways to make their embodied ideas into a meaningful representation of the definition.

Figure 7 (overleaf, taken from Tall, (2001)) shows how formal proof interrelates with thought experiments based on the concept image, with the movement being in *both* directions. Imagery can be used to suggest theorems to prove. On the other hand, amongst theorems that might be proved, there are some that are more enlightening than others. These are called *structure theorems*. A structure theorem is deduced by formal methods from the concept definition. However, it has a special function: it formulates the *structure* of an axiomatic system in such a way that it may be used as a richer image on which new thought experiments may be performed. Such structure theorems exist throughout formal mathematics. For instance, a finite dimensional vector space over a field F defined axiomatically must have the structure of *n-tuples* in the field F. Or a group defined axiomatically must, by Cayley's Theorem, have the structure of a group of permutations of a set. Or that a complete ordered field has the structure of the real number line.

Structure theorems are important because they relate axiomatic concepts to mental embodiments that confirm the nature of the structure. This not only allows the individual to use the structure as a mental embodiment to think, it allows an interplay between imagery, often based on embodiment, and formal structure, based on logic. Lakoff and his colleagues (Lakoff & Johnson, 1999, Lakoff & Nunez, 2000) claim that *all* mathematics comes from embodiment. It certainly underpins conceptual growth, as we have seen. However, what is certainly true is that formal mathematics attempts to base its deductions on more than just imagistic thought experiments, so that the theorems that are proved work not only in a single embodiment (such as that of the geometric figures in Euclidean geometry) but in *any* structure that obeys the axioms.

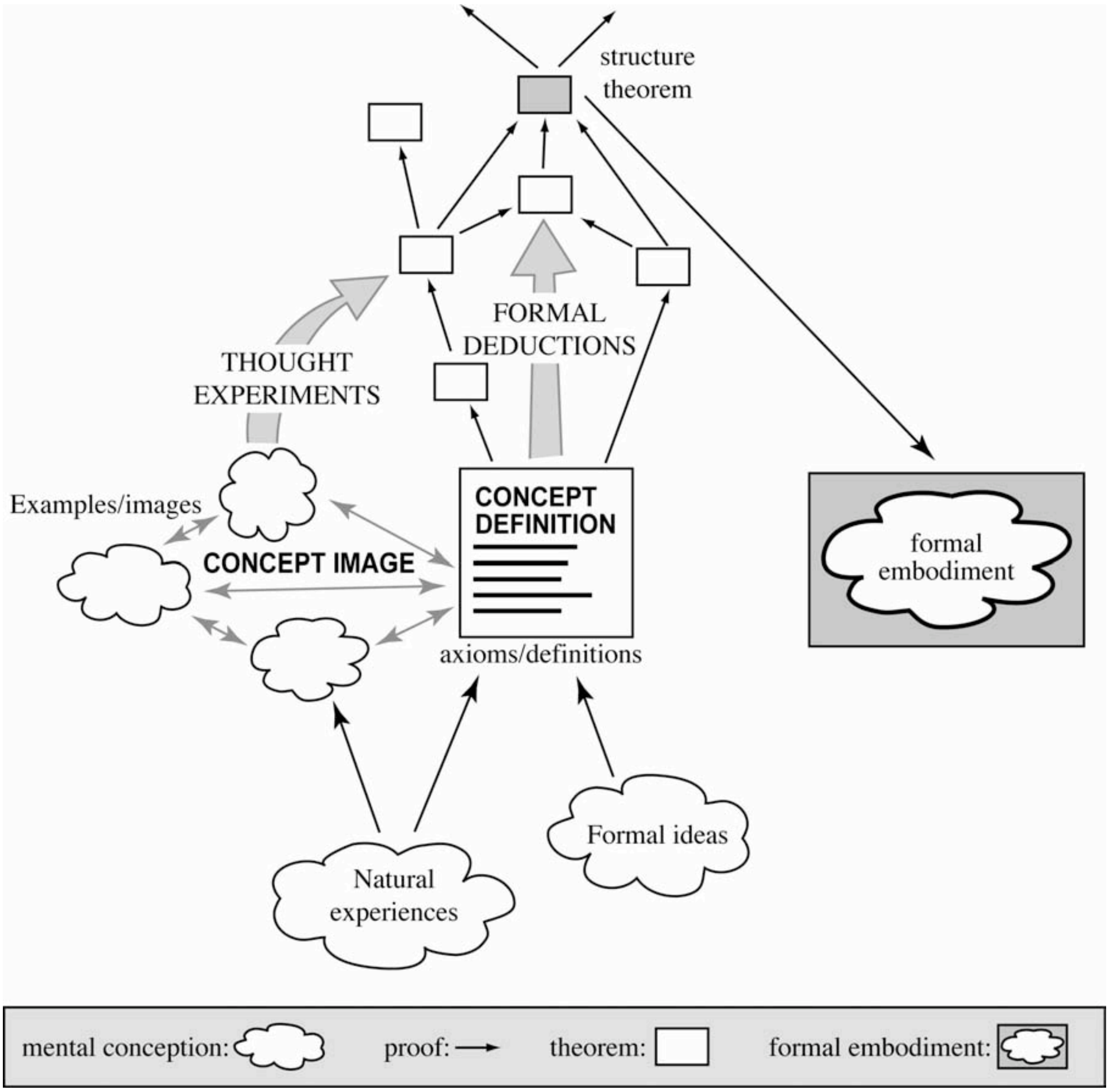

Figure 7: Interplay between concept image and formal proof

This is the power of formal proof, over and above the power of insight that may come from a particular embodiment. It is the fundamental reason why formal proof is the foundation of the work of research mathematicians. However, such formal proof is the pinnacle of mathematical development, and requires the implicit understanding shared by mathematicians to sustain it. To reach such a refined level requires a long and complex cognitive development. As this development progresses, the different warrants for truth in the different worlds of mathematics develop different meanings through different ways of building knowledge. Without the biological development of the human brain we would not have the power of mathematics. We would do well to remember this and work to understand the nature of this development and how we may use this knowledge in the educational growth of our children.

References

Davis, P. J. & Hersh, R. (1981). *The Mathematical Experience*. Birkhauser: Boston.

Freudenthal, H. (1983). *Didactic phenomenology of mathematical structures*, Dordrecht: Reidel.

Gray, E. M. & Tall, D. O. (1994). Duality, ambiguity and flexibility: A proceptual view of simple arithmetic. *Journal for Research in Mathematics Education*, 25, 2, 115–141.

Chin, E.-T. (2002). *Building and using concepts of equivalence class and partition.* PhD Thesis, University of Warwick.

Chin, E.-T. & Tall, D. O. (2002). Mathematical Proof as Formal Procept in Advanced Mathematical Thinking. This volume.

Guttierez, A., Jaime, A. & Fortuny, J. M. (1991). An alternative paradigm to evaluate the acquisition of the van Hiele levels. *Journal for Research in Mathematics Education*, 22, 237–251.

Johnson (ed), (1989): *Children's Mathematical Frameworks 8–13: A Study of Classroom Teaching*. Windsor UK: NFER-Nelson.

Lakoff, G. & Johnson, M. (1999). *Philosophy in the Flesh.* New York: Basic Books.

Lakoff, G. & Nunez, R. (2000). *Where Mathematics Comes From.* New York: Basic Books.

Pinto, M. (1998). *Students' Understanding of Real Analysis.* PhD Thesis, Warwick University.

Rodd, M. M. (2000). 'On mathematical warrants', *Mathematical Thinking and Learning*, Vol. 2 no. 3, 221-244.

Rodd, M.M. (1997). Beliefs and their warrants in mathematics learning, in Pehkonen, E. (Ed.) *Proceedings of the 21st Conference of the International Group for the Psychology of Mathematics Education*, Vol.4, pp.64-71. Lahti, Finland.

Rodd, M.M. (1998). *Mathematical warrants, objects and actions in higher school mathematics*. Ph.D. thesis, the Open University.

Rosch, E., Mervis, C., Gray, W., Johnson, D., Boyes-Braem, P. (1976). Basic Objects in Natural Categories. *Cognitive Psychology* 8: 382–439.

Stewart Ian & Tall, David, (1977). *Foundations*, Oxford University Press.

Stewart Ian & Tall, David (2001). *Algebraic Number Theory and Fermat's Last Theorem.* Peters.

Tall, D. O. (ed.) 1991. *Advanced Mathematical Thinking*. Kluwer: Dordrecht.

Tall, D. O. (1999). The Cognitive Development of Proof: Is Mathematical Proof For All or For Some? In Z. Usiskin (Ed.), *Developments in School Mathematics Education Around the World*, vol. 4, 117–136.Reston, Virginia: NCTM.

Tall, D. O., (1995), Cognitive growth in elementary and advanced mathematical thinking. In D. Carraher and L. Miera (Eds.), *Proceedings of PME X1X*, Recife: Brazil. Vol. 1, 61–75.

Tall, D.O., Gray, E., M., Bin Ali, M., Crowley, L., DeMarois, P., McGowen, M., Pitta, D., Pinto, M., Thomas, M., Yusof, Y., (2001). Symbols and the Bifurcation between Procedural and Conceptual Thinking, *The Canadian Journal of Science, Mathematics and Technology Education*, 1, 80–104.

Thurston, W. P. (1994). On proof and progress in mathematics, *Bulletin of the American Mathematical Society*, **30** 2, 161–177.

Van Hiele, P. M. (1986). *Structure and Insight. A Theory of Mathematics Education.* London: Academic Press Inc.

Appendix

David Tall: Selected Papers on Proof

(downloadable from www.davidtall.www/papers)

1989a The nature of mathematical proof, *Mathematics Teaching*, 127, 28-32.
A humorous reflection on the nature of proof written for teachers and teenagers, with some deeper considerations.

1988h (with John Mills) From the Visual to the Logical, *Bulletin of the I.M.A.* 24 11/12 Nov–Dec, 176–183.
Thinking about the relationship between visual intuition (supported by computers) and logical proof.

1992e The Transition to Advanced Mathematical Thinking: Functions, Limits, Infinity, and Proof, in Grouws D.A. (ed.) *Handbook of Research on MathematicsTeaching and Learning*, Macmillan, New York, 495–511.
A review of the literature concerning the transition to formal mathematics, including a consideration of the transition to formal proof.

1992m Construction of Objects through Definition and Proof, *PME Working Group on AMT*, Durham, NH.
Preliminary ideas concerning the way in which mathematical concepts are *constructed* from definitions using formal proof.

1995 Cognitive growth in elementary and advanced mathematical thinking. In D. Carraher and L. Miera (Eds.), *Proceedings of PME X1X*, Recife: Brazil. Vol. 1, 61–75.
The first publication of ideas that led to 'three worlds of mathematics.'

1995f Cognitive Development, Representations & Proof. In *Justifying and Proving in School Mathematic*s, Institute of Education, London, 27-38.
A first consideration of the different types of proof available depending on the representations that are meaningful for the learner.

1996b A Versatile Theory of Visualisation and Symbolisation in Mathematics, Plenary Presentation, *Proceedings of the 46th Conference of CIEAEM*, Toulouse, France (July, 1994), 1, 15–27.
The first presentation on ideas that became 'the three worlds of mathematics', given in July 1994.

1997d (with Tony Barnard), Cognitive Units, Connections and Mathematical Proof, *Proceedings of PME 21*, Finland, 2, 41-48.
My first attempt to develop Tony Barnard's notion of 'cognitive unit' relating it to the proof that root 2 is irrational.

1998a (with Liz Bills), Operable Definitions in Advanced Mathematics: The case of the Least Upper Bound, *Proceedings of PME 22, Stellenbosch, South Africa*, 2, 104-111.
A paper which I regard as being seminally important, (although I have not followed it up as carefully as it deserves). This paper investigates the development of students' ability to use a definition in an operable manner to deduce theorems as they follow a lecture course introducing a definition and then using it to build a formal theory. Of the five students interviewed over two 10-week terms, none were able to make the definitions operable. One struggled with the definition, one learned it because of the interviews, the other three simply built intuitively on their concept imagery.

1998b (with Adrian Simpson), Computers and the Link between Intuition and Formalism. In *Proceedings of the Tenth Annual International Conference on Technology in Collegiate Mathematics*. Addison-Wesley Longman. pp. 417-421.
A further development of earlier ideas on visual intuition and formal logic.
1999j The Cognitive Development of Proof: Is Mathematical Proof For All or For Some? In Z. Usiskin (Ed.), *Developments in School Mathematics Education Around the World*, vol. 4, 117-136. Reston, Virginia: NCTM.
A development of 1995f (above) considering how child development leads to a sequence of different stages of proof development related to the development of enactive, visual, symbolic, formal modes of thought.
1999a The Chasm between Thought Experiment and Mathematical Proof. In G. Kadunz, G. Ossimitz. W. Peschek, E. Schneider, B. Winkelmann (Eds.), Mathematische Bildung und neue Technologien, Teubner, Stuttgart, 319-343.
A plenary lecture expanding on the transition from elementary to advanced mathematical thinking, including a discussion of the work of Marcia Pinto. (Published before 1999j, but given after.)
1999g (with Márcia Maria Fusaro Pinto), Student constructions of formal theory: giving and extracting meaning. In O. Zaslavsky (Ed.), *Proceedings of the 23rd Conference of PME, Haifa, Israel,* 4, 65–73.The emergence of the ideas of 'natural' and 'formal' thinking, based on a grounded theory built from observations of first year university students coping with an analysis course.
2001j (with Marcia Maria Fusaro Pinto) Following student's development in a traditional university classroom. In Marja van den Heuvel-Panhuizen (Ed.) *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education* 4, 57-64. Utrecht, The Netherlands.
A development of 1999g (above) from the PhD thesis of Marcia Pinto, based on the distinction between natural and formal thinking.
2000c (with Ehr-Tsung Chin), Making, Having and Compressing Formal Mathematical Concepts. In T. Nakahara and M. Koyama (eds) *Proceedings of the 24th Conference of the International Group for the Psychology of Mathematics Education 2*, 177-184. Hiroshima, Japan.
A preliminary study of the development of equivalence relation and partition, from intuitive ideas through definitions, deductions and theorems
2000 Proof and the Transition from Elementary to Advanced Mathematical Thinking, *Proceedings of the Conference on Secondary Mathematics Teaching*, Athens, May 12–14, 2000 (to appear).
A plenary lecture developing my theory of transition from elementary to advanced axiomatic thinking.
2001d (with Tony Barnard) A Comparative Study of Cognitive Units in Mathematical Thinking. In Marja van den Heuvel-Panhuizen (Ed.) *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education* 2, 89–96. Utrecht, The Netherlands.
A preliminary version of a longer paper on Cognitive Units, Connections and Compression in Mathematical Thinking. This considers the cognitive constructions and links involved in mathematical thinking and proof.
2001g (with (Abe) Ehr-Tsung Chin) Developing Formal Mathematical Concepts over Time. In Marja van den Heuvel-Panhuizen (Ed.) *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education* 2, 241-248. Utrecht, The Netherlands.
An analysis of the relationship between the formal definitions of equivalence relation and partition, with the equivalence relation being more related to a formal definition and the partition (which is logically equivalent) being cognitively associated with intuitive visual ideas.
2001p Natural and Formal Infinities. *Educational Studies in Mathematics*, 48 (2&3), 199–238.
This paper considers the difference between cognitive and formal conceptions of infinity in great detail, with particular reference to how formal thinking can lead to structure theorems that have natural interpretations in visual form. It therefore has wider implications for the development of formal proof.
2002a (with Marcia Pinto), Building formal mathematics on visual imagery: a theory and a case study. *For the Learning of Mathematics*. 22 (1), 2–10.
A focus on a natural learner building formalism from coherently organised visual intuition.