The Cognitive Development of Proof:
Is Mathematical Proof For All or For Some?

David Tall

Mathematics Education Research Centre
Institute of Education
University of Warwick
Coventry CV4 7AL
e-mail: David.Tall@warwick.ac.uk

Proof is considered a foundational topic in mathematics. Yet, it is often difficult to teach. In this paper, I suggest that different forms of proof are appropriate in different contexts, dependent on the particular forms of representation available to the individual, and that these forms become available at different stages of cognitive development. For a young child, proof may be by way of a physical demonstration, long before sophisticated use of the verbal proofs of euclidean geometry can be introduced successfully to a subset of the school population. Later still, formal proof from axioms involves even greater difficulties that make it appropriate for a few, but impenetrable to many. At this formal stage of development, I will identify two different strategies that students adopt to come to terms with formal definition and deduction. Either strategy may be successful, but both are cognitively demanding and prove difficult for many to achieve. This leads to the observation that formal proof is appropriate only for some, that some forms of proof may be appropriate for more, and that, if one allows the simpler representations of proof such as those using physical demonstrations, perhaps some forms of proof are appropriate for (almost) all.

Introduction: What is "Proof"

There is an old English Music Hall joke that goes something like this:

> Comedian**: I can prove that I am not here**.
> Straight Man: *Go on, then.*
> **Well, I'm not in Rome am I?**
> *No.*
> **And I'm not in Paris?**
> *No.*
> **Well, if I'm not in Rome and I'm not in Paris, then I must be somewhere else.**
> *Yes.*
> **And if I'm somewhere else, I can't be here!**

Formal proof often feels like this to students. Despite the fact that each step seems to follow from the one before, in total it makes no sense. In this case, the

This paper was presented at the Conference of the University of Chicago School Mathematics Project, August, 1998.

argument fails because of a subtle change in the meaning of being "somewhere else". This "proof" almost seems to work, line by line, but it fails as a whole.

Although some individuals may be satisfied when they can follow a proof line by line, others require more, that it conveys insight:

> To understand the demonstration of a theorem, is that to examine successively each of the syllogisms composing it and to ascertain its correctness, its conformity to the rules of the game? ... For some, yes; when they have done this, they will say: I understand. For the majority, no. Almost all are much more exacting; they wish to know not merely whether all the syllogisms of a demonstrations are correct, but why they link together in this order rather than another. In so far as to them they seem engendered by caprice and not by an intelligence always conscious of the end to be attained, they do not believe that they understand. (Poincaré, 1913, p. 431)

The process of mathematical proof therefore has two different purposes. One is to show that an assumption leads to a stated conclusion in a sequence of logical steps. The other is to give meaningful insight as to how and why the conclusion follows from the assumptions. Note that these two can be independent, that is, a logical proof need not be meaningful and a meaningful proof need not be logical. The latter, giving insight that convinces, is a form of proof that can be help learners make connections to support their thinking.

Proof also has a further purpose in which the results of the theorems proved are used to build up a systematic theory. Proof, therefore, has the following contrasting features:

(1) (*Local*)
   - (a) (*logical*): based on explicit assumptions, a proof is used to deduce step by step that certain consequences follow,
   - (b) (*meaningful*) providing insight as to how and why the consequences follow from the given assumptions.

(2) (*Global*) such consequences can then be used as building blocks (or "relay results" in the words of Hadamard, 1945) to construct a systematic mathematical theory.

The form of proof traditionally presented in school is (or was) Euclidean geometry, intended as an introduction to both (1) (in a meaningful sense) and (2). However, empirical research in the USA by Senk (1985) has showed that only 30% of students in a full-year geometry course reached a 70% mastery on a selection of six problems in Euclidean proof. In other words, *traditional Euclidean proof may be suitable for some, but not for all.*

The NCTM standards in the USA proposed that there should be increased attention on short sequences of theorems and decreased attention to Euclidean geometry as an axiomatic system, thus favoring (1) (particularly in the meaningful version (1)(b)) over (2). In England the demise of geometry has proceeded even further, being replaced in the English National Curriculum by the study of “Shape and Space”. The curriculum for 5 to 16 year-olds has only two explicit references to Euclidean proof, both at the top level. “Shape and space” mentions:

> … knowing and using angle and tangent properties of circles,

and “Using and Applying Mathematics” indicates that pupils should:

> … find their own proof that the angle in a semi-circle is a right angle and its converse, stating what prior results have been assumed.

Euclidean Geometry as a structured theory in the British National Curriculum is dead. “Proof” remains in the curriculum, but it is now in terms of justifying conjectures, often using generalized arithmetic or algebra. This is something that I regret, and yet, if one looks at the cognitive development of the individual, one finds that “proof” is not a single all-embracing methodology exemplified by Euclidean proof. As the child develops, different contexts are perceived in different ways, each having its own form of justification. In such a context, the form of proof used in school is intended to be a meaningful activity (1b) rather than a logical deduction (1a). In the remainder of this paper we will look at ways in which proof can be presented meaningfully as the child develops and close with a consideration of proof as a systematic theory taught at university level.

Cognitive development of representations and proof

The cognitive development of the child is characterized by the construction of increasingly sophisticated mental concepts. Bruner (1966) formulated the development in terms of three kinds of representation: enactive, iconic and symbolic. The most basic form of communication is *enactive*, using gestures and physical actions to convey ideas. The next is *iconic*, using pictures or diagrams as physical representations. The third is *symbolic*, by which Bruner (1966, pp. 18,19) meant not only “language in its natural form” but also the two “artificial languages of number and logic.”

My own interpretation (Tall, 1995) sees the symbolism referred to by Bruner acting in different ways. Natural language is used to describe visual images and to explain the operations of number and logic. I therefore see language playing an overall role in the development of the other representations. In operating on the real world, we perceive things and act upon them. These distinct aspects lead to quite different types of cognitive development. *Perception of* objects leads to physical exploration and verbal description exploring shape and space, only later developing into the verbal proof of Euclidean geometry. *Action on* objects leads to sorting, counting, ordering, typical of number and arithmetic. In elementary arithmetic, proof occurs mainly in terms of generalized arithmetic using algebraic notation and manipulation.

Formal logic is an altogether more subtle conception which I see as building on experiences of the world and only becoming operative when the properties involved can be specified and used in a deductive manner. A fully formal axiomatic treatment is developed at a more sophisticated level. This cognitive development of mathematical knowledge is shown in figure 1.

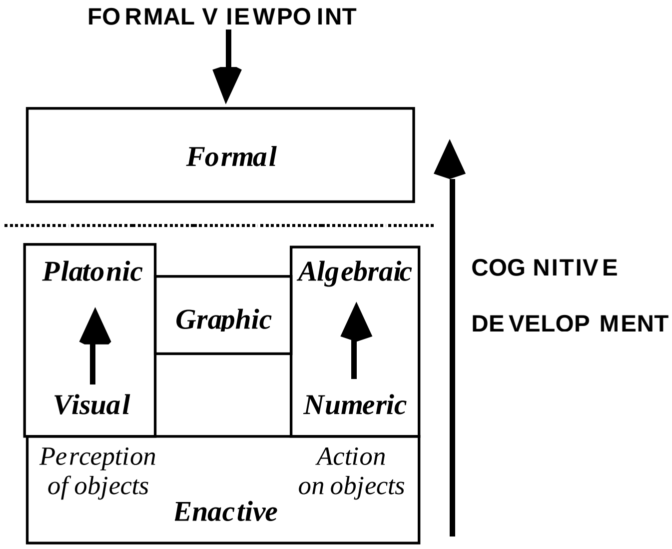


Figure 1: Cognitive development of representations

The visual development begins with a focus on *perception of objects* in the environment which are then classified using verbal descriptions (such as points, lines, circles, squares, etc., in plane geometry). Language enables us to refine our ideas from *visual* perception of physical objects to the imagination of *platonic* objects in a cognitive development broadly formulated by Van Hiele (1959, 1986). A perceived straight line, drawn with pen on paper, can be imagined to represent a theoretical straight line, perfectly straight, with no thickness, capable of arbitrary extension in either direction[1]. Proof in geometry can involve physical demonstrations in the initial stages, moving on to the use of verbal definitions and platonic imagery in Euclidean geometry.

On the other hand, the actions we perform in counting, using the language of number, become represented by number symbols enabling most of us to carry out the process of counting and also talk about the concept of number. Gray & Tall, (1994) describe a theory of symbol-use where a symbol that dually functions to evoke process or concept is called a *procept.* Such symbols allow individuals both to *do* mathematics and to *think* about it, starting with counting and arithmetic, and then (for some) on to the manipulation of algebra. Throughout the world, various parts of this development prove difficult for some, for instance in the arithmetic of fractions and the introduction of algebra. There is no research that I know of which claims to make algebra available for all. Proof in arithmetic and algebra usually involves either checking through

[1] Many individuals imagine a “point” to be an “arbitrarily small” mark, or a line to have “arbitrarily small” thickness. This leads to beliefs about “infinitesimal quantities” on the number line giving a mental model different from that of the real numbers. These further complications are discussed in Tall (1980), Cornu (1991).

carrying out calculations or confirming general statements using algebraic formulations and manipulations.

Straddling the visual/platonic and the numeric/algebraic developments are links between the visual and symbolic through visual representations of numerical relationships or the use of the real line and the coordinate plane to visualize symbolic relationships. There are then parallels between algebraic manipulation in coordinate geometry and proofs in Euclidean geometry.

At the top of the developmental diagram is formal (axiomatic) proof. I intend to show later that there is a significant cognitive barrier in the transition to formal proof that causes great difficulty for many students (Tall, 1992). Its existence also represents a considerable barrier for us as mathematicians attempting to plan a curriculum for learners. We see mathematical structure from an expert viewpoint, from a position attained after a lifetime's development. To build a curriculum for learners requires us to attempt to see the ideas within a cognitive growth, not only in terms of the various representations that become available but also taking account of the wide variations between individuals. We begin this journey by looking at different kinds of proof which become available using different representations. Here we strive for meaningful proof (1b) rather than formally logical proof (1a).

Enactive proof

At the most primitive level, *enactive proof* involves carrying out a physical action to demonstrate the truth of something. This invariably involves visual and verbal support, but the essential factor is the need for physical movement to show the required relationships. For instance, I was privileged to see an imaginative teacher show her class of five-year-old children that "three and two is the same as two and three". She did this using some beads in her necklace, separating two in one hand and three in the other, switching it round so that the arrangement could be seen as two and three or three and two, giving the same number five each time.

Enactive proof can also be used in more sophisticated contexts. For instance, to show that a triangle with equal sides has equal angles, one might cut out a typical triangle made of paper and fold it down its axis of symmetry to show that when the two equal sides match, so do the base angles.

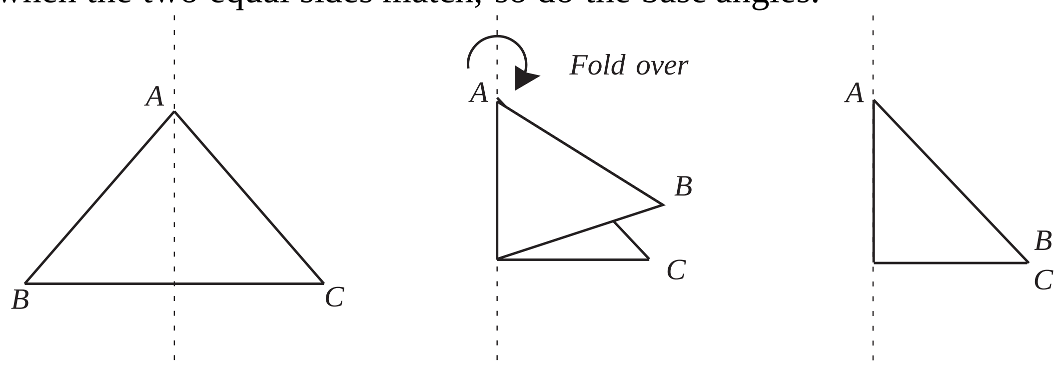


Figure 2: An enactive (visual) proof that equal sides imply equal angles

Such a proof involves specific examples seen as representative *prototypes* of a class of examples.

Visual (Enactive) Proof of Geometric Statements

Visual proof often involves enactive elements (and usually has verbal support). For instance, the famous classical Indian proof of Pythagoras takes four copies of a right angled triangle with sides a,b and hypotenuse c, and places them in two different ways in a square side a+b. The remaining area can be expressed as two squares area a^2 and b^2, or as a single square area c^2, giving $a^2 + b^2 = c^2$.

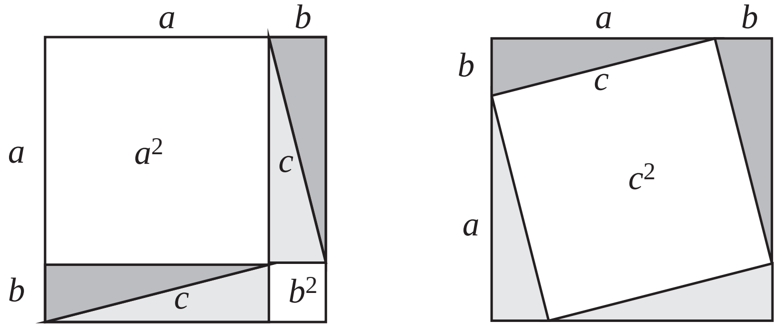


Figure 3: An (enactive) visual proof of Pythagoras (after Bhaskara)

To "see" this proof, it is essential to be able to imagine how the triangles can be moved around from one configuration to another. This movement is not obviously translatable into a formal sequence of logical steps to give the standard Euclidean proof (or something equivalent). Thus it happens that a visual proof might be meaningful in its own terms without necessarily leading straight into a formal proof.

Note also that any actual drawing will have specific (positive) values for a and b, but such a diagram can be seen as a *prototype*, typical of *any* right-angled triangle. This gives a kind of proof which is often termed "generic"; it involves "seeing the general in the specific".

Graphic Proof of Numeric and Algebraic Statements

The idea underlying certain arithmetical statements can be "seen" to be true by using visual configurations in a generic way as prototypes. For instance, a picture of a 2×3 array can be seen as 2 rows with 3 in each row or 3 columns with 2 in each column (figure 4).

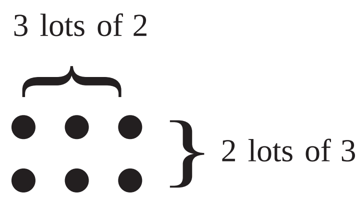


Figure 4: a graphic proof that 3×2 is the same as 2×3

Such a representation is called "graphic" (relating back to figure 1) because it links symbols and pictures. The "proof" occurs by seeing the *same* diagram in two different ways (as rows or as columns). This is less dependent on an enactive rearrangement and more dependent on re-focusing attention to see the array as rows or columns. It may be also seen as being typical of a class of similar pictures, such as 4×5 or 27×13, each a typical prototype for the *general* statement

$$m \times n = n \times m$$

for whole numbers m, n.

Likewise, the algebraic identity $a^2 - b^2 = (a+b)(a-b)$ can be visualized in the generic diagram given in figure 5.

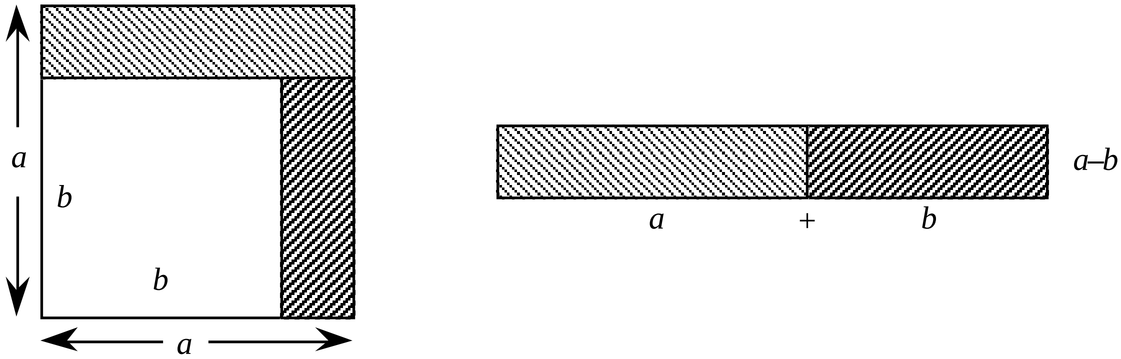


Figure 5: Taking a square side b from a square side *a* and rearranging what is left as $(a-b)\times(a+b)$

This graphic proof again has enactive elements to see the dynamic rearrangement of the parts. The diagram given applies only to *positive* values of *a* and *b*. Nunn (1914) represented positive and negative in different colors, using a variety of diagrams to show all possible cases, but their complexity seems likely to cause them to outlive their usefulness. (If you don't believe this, then try to draw pictures for *a*, *b* positive or negative. Symbolism has the power that *one* equation represents all cases, pictures often need to consider various cases separately.)

Proof in arithmetic by specific and generic calculation

Arithmetic, as a computational activity, usually has little proof involved, other than the *checking* of calculations perhaps saying something like 24532 times 34513 cannot equal 846672915 because the units digit is not even.

Proof in a more general sense occurs when the quantifiers "all" or "some" or "none" are introduced. For instance, the statement that $\sqrt{2}$ is irrational equivalently says that the square of any rational number cannot be 2. Here a *generic* proof is possible, namely that any specific number one attempts to square cannot give the value 2. First one notes that any fraction in lowest terms can be factorized into primes, e.g.

$$\frac{9}{40} = \frac{3^2}{2^3 \times 5}$$

and the squaring of this number *doubles* the number of each prime factor to give

$$\left(\frac{9}{40}\right)^2 = \frac{3^2}{2^3 \times 5} \times \frac{3^2}{2^3 \times 5} = \frac{3^4}{2^6 \times 5^2}$$

so the primes occurring in the factorization of numerator and denominator of a square number all occur an even number of times. Hence the square of any fraction cannot equal 2 which factorizes as 2/1 and has an *odd* number of 2's in the numerator.

In Tall (1979), I showed that students in the first year of university expressed a strong preference for the generic proof over the standard proof by contradiction. I also remarked that this does *not* mean that the generic proof is preferable in the long term. Proof by contradiction is an essential element in formal mathematics and needs to be addressed, even though it involves significant cognitive difficulties. These difficulties are more subtle than is often assumed.

Barnard & Tall (1997) investigated students understanding of the standard proof by contradiction that √2 is irrational. It transpires that it is nested with a contradiction proof inside a contradiction proof. For, having considered the possibility that $\sqrt{2} = a/b$ and squaring to get $a^2=2b^2$, it is necessary to show

a^2 is even implies a is even.

Students often step over this deduction without fully understanding it, for instance Student S responded asserting the authority of the lecturer: "the root of an even number is even—he just assumed it." When the interviewer challenged this statement with an example, the following conversation took place:

Interviewer:	So the root of six is even.
Student S:	Good point. [five seconds pause]
Interviewer:	If a number is not even, what is it?
Student S:	It's odd.
Interviewer:	So you've got a choice of odd or even, does that help you?
Student S:	Yeah, I see, it's got to be rational, I think, so … a rational root is either … odd or even and if the square is even, then the rational root is even. Is that clear?
Interviewer:	Uh, well …
Student S:	So what I'm thinking is the root of 4, 4's even and 2's even, root of 16 equals 4, … 's even. I can't remember any other simple squares in my head that are even …

The simple proof by contradiction that "If a^2 is even then a is even" is difficult for Student S; he seems to see a number and its square being *simultaneously* even or odd, not that one is the *consequence* of the other. It is ironic that the one proof we give to students as an archetypal example of proof by contradiction should contain such hidden difficulties.

Algebraic proof by algebraic manipulation

Algebra has the ability to express arithmetic ideas in a *general* notation and so has more scope for proof than generic arithmetic. For instance, the fact that the sum of two consecutive odd numbers is a multiple of 4 may be expressed algebraically by noting that $2n+1$ plus $2n+3$ is $4n+4$. Such a proof is carried out by using a suitable algebraic representation and performing an algebraic manipulation (in this case the addition of two expressions).

This is the most commonly occurring method of "proof" in the English National Curriculum, and occurs widely in numerical investigations. It uses algebraic manipulation rather than logical deduction. In general, logic has a low (almost non-existent) priority in the British National Curriculum, although it still has a presence in the curricula of other countries (for instance, in Italy).

Some facts in algebra (such as $a+b=b+a$) are believed to be true from arithmetic and other evidence (such as visualizing alternative layouts for numbers of items or different lengths). Others are "proved" by symbolic manipulation, such as the identity

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

being proved by multiplying out the brackets on the left-hand side and canceling the terms ba and $-ab$. However, it is not easy to make plain what should be assumed "true" and what needs to be proved at any given stage. For instance, which of the following statements are "evidently" true and which are false?

$$a+b = b+a \text{ .... } (i)$$
$$a(b+c) = ab+ac \text{ .... } (ii)$$
$$(a+b)(a-b) = a^2-b^2 \text{ .... } (iii)$$

Theoretically, we might suggest that (*i*) and (*ii*) are evidently true and (*iii*) requires proof. But this relies on our mathematical experience of the axioms of a number system. The learner is placed in a very different position. His or her experience in arithmetic may be that (*i*) has long been "known" to be true as "the commutative rule". Equation *(ii*) might be demonstrated in a number of ways (by generic arithmetic, or a picture). Item *(iii)* can equally be illustrated by generic arithmetic examples or by a picture (figure 5). It is therefore not obvious to the learner which of these is "known" to be true and which needs to be "proved". Thus, although algebraic notation may be used to formulate a general arithmetical pattern to "prove" a general statement, the manipulation of algebra must await the development of axiomatic proof to specify some "truths" as axioms, and then deduce the others.

Euclidean Proof as a verbal translation of generic visual proof

Euclidean proof is often seen as being a good starting point to develop the rigor of logical proof. As proposed in the books of Euclid it seems to have the form of a major systematic theory. However, as encountered in school, individual proofs are almost always verbal translations of what is seen in a visual picture involving a certain geometric configuration. Consider figure 6, which accompanies the theorem that if ΔABC has $AB=AC$ then $\angle B=\angle C$.

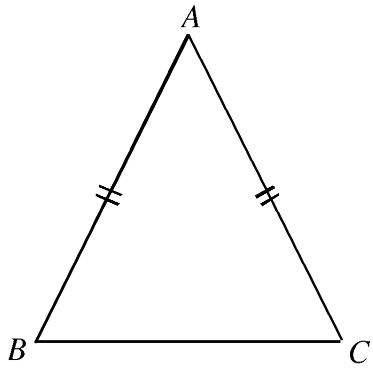


Figure 6 a specific isosceles triangle

The verbal proof applies not just to the specific picture drawn, but generically to the whole class of figures represented by the theorem. For instance, the proof that applies not just to this triangle *ABC*, but to all the triangles in figure 7.

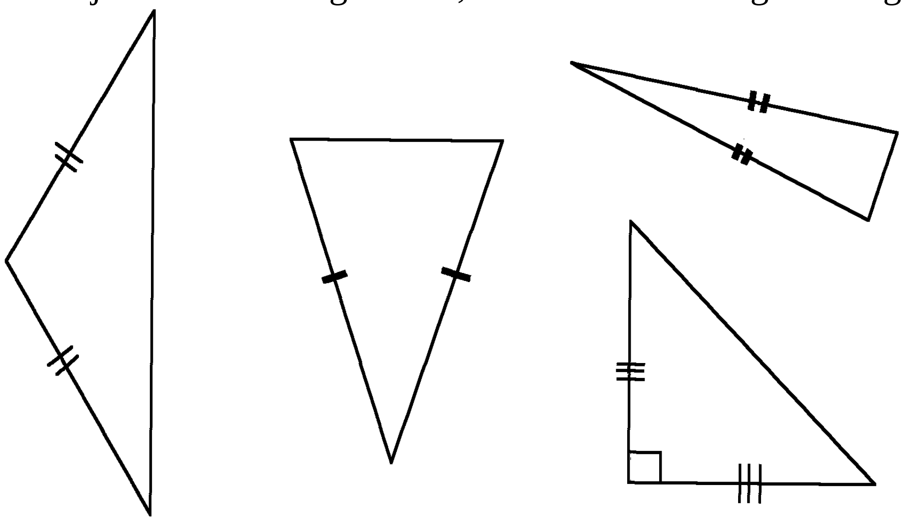


figure 7 : other isosceles triangles

In this way, Euclidean proof is *verbal generic proof* applying to the whole class of geometric figures having the given properties. It is interesting to note that the Greeks used to regard a Euclidean proof first as a proof for the *particular triangle* drawn on the page, then they emphasized that the "same" proof would work in all other cases.

Weaknesses in Visual Proof

In the nineteenth century, it was realized that the verbal language of Euclidean geometry contained implicit beliefs which were not part of the formal definitions. For instance, the idea that the diagonals of a rhombus meet "inside" the figure, where "inside" had not been defined in the list of axioms and common notions. This proved a shock to the system which was made worse when the functions of mathematical analysis proved to have seemingly unbelievable properties (such as the existence of functions continuous everywhere but differentiable nowhere). Visual ideas became suspect and untrustworthy, despite the manner in which they often seem so convincing.

I suggest that the fundamental problem lies in the nature of the visual representation used in the proof. As it is a *prototype* for the proof, its applicability only extends to the class of examples for which it is prototypical. Thus the generic proof of $m \times n = n \times m$ given earlier applies in the given pictorial form only to positive whole numbers and the visual proof of the algebraic identity for the difference of two squares applies initially only to positive real numbers.

As concepts change in meaning—from enactive, through visual or symbolic, and on to formal—different kinds of proof may convince the individual. But what is satisfactory to an individual at one stage of development can (and often *does*) prove to be unsatisfactory later.

An archetypal example of this is the proof of the intermediate value theorem, that a function which is negative at *a* and positive at *b* and continuous from *a* to *b* must be zero somewhere between *a* and *b*.

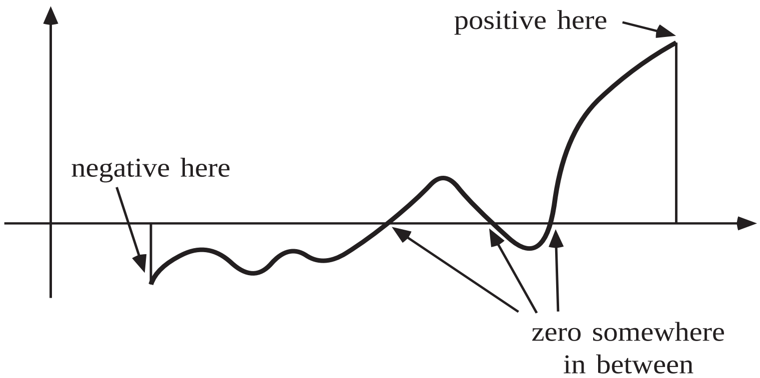


Figure 8: The Intermediate Value Theorem

Enactively the notion of "continuous function" is something that is drawn "continuously" without taking the pencil from the paper. The theorem is "evidently true" in this context because if one attempts to draw such a graph then it *must* cross the axis at least once in between.

Visually, once the graph is drawn as a static object, the notion of continuity becomes the idea of being "all in one piece", which corresponds to the notion of "connected". Again, even when viewed as a static picture, (as above), one cannot but imagine a point moving along the graph and see evidently that the theorem "must" be true.

The need for formal proof

The previous instance of the Intermediate Value Theorem seems to be true both *enactively* (in drawing a graph) and *visually* (in seeing a graph "all in one piece"). Yet, is it true in every possible case?

Consider the function $f(x)=x^2-2$. This is negative for $x=1$, positive for $x=2$ so, according to the theorem, there is a value in between where $f(x)=0$. There certainly is, it is $\sqrt{2}$. But here is the problem. The number $\sqrt{2}$ is *real, but not rational*. Suppose we consider the function to be defined only for *rational values*. We again get a function f defined everywhere on the rationals where $f(1)$ is negative and $f(2)$ is positive. But there is no *rational* number x between 1 and 2 such that $f(x)=0$. The theorem in this case is *not true*. Worse still, we cannot visually distinguish between rational and irrational numbers on a physical picture of the number line, although we may imagine the difference in our mind's eye.

To overcome the difficulty we must move on from physical pictures and beyond mental pictures of lines that fail to distinguish between rationals and irrationals and on to precise definitions and deductions. For instance in formal mathematics, the real number system $\boldsymbol{R}$ is defined as a "complete ordered field" and a "continuous function" is defined in terms of the ε-δ definition. This allows the Intermediate Value Theorem to be stated correctly in terms of a function that is continuous on a real interval $a \le x \le b$.

This change from elementary mathematics to formal proof involves a huge cognitive struggle. Whereas in elementary mathematics the properties of numbers are developed through practical experience, in formal mathematics we must start again from scratch, selecting specific properties as axioms and deducing everything from them. This means *everything*, including such "obvious" facts as 2+2=4 should be formally proved. In practice, we *never* prove *everything*. We do a few such deductions, then assume that we could do any others required. The expert sees this as being economical, but the learner initially sees it as confusing. Which facts does one "know" and which are "to be proved"?

Sometimes theorems that seem enactively or visually "obvious" are false in a more general context. On the other hand, theorems that are "obvious" in an informal sense may also be formally true, but the formal proof is not obvious. An example is the "Jordan curve theorem" (that every closed path in the plane that does not cross itself divides the plane into two regions, the "inside" and "outside").

Exacerbating the situation is the great complexity of quantifiers that occurs in definitions and deduction in formal mathematics (such as those concerning limits and continuity in analysis). In such cases, the individual often has to struggle to "follow" a proof in the first place before being convinced that such a proof is acceptable. Often the battle with the proof fails and the student may resort to rote-learning for reproduction in an exam with little understanding of how or why the proof works.

Empirical studies of proof at university

Over recent years I have taken to tracking the development of individual students through a specific course in mathematics, attempting to divine the changes in their competencies and belief structures. In particular, how do students cope with definitions and proofs? Many years ago, 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.
> (Tall & Vinner 1981, p.152)

On the other hand:

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

The way in which concept definitions can be used in formal proofs is very much intertwined with all the other concept imagery that the individual has at the time. The important factor is to be able to distinguish between this broad concept image and the subset consisting of the *formal image*, consisting *only* of those things that can be deduced from the formal definition.

In Bills & Tall (1998), an *operable* definition was formulated as:

> A (mathematical) definition or theorem is said to be *formally operable* for a given individual if that individual is able to use it in creating or (meaningfully) reproducing a formal argument.

We found that it was not a simple matter of *learning* the definition and then *using* it to develop a formal theory. Bills & Tall followed the use of the notion of “least upper bound” through twenty weeks of an analysis course. Only one student of those studied learned the definition. He was able to go on to give a good explanation of the Riemann integral for a continuous function in terms of greatest lower bound of upper sums and least upper bound of lower sums. On further questioning it became clear that he did this without knowing the definition of a continuous function, instead relying on his concept imagery of a continuous function as a curve whose graph can be drawn without taking the pencil from the paper. What *appeared* to be a formal discussion of Riemann integral was based in part on a *formal* definition of least upper bound and in part on an *informal* notion of continuity.

Another student, who was generally more successful, did not learn the definition, and could not reproduce it exactly as she was studying. Despite this, throughout the course she always seemed to be able to discuss the ideas in a coherent and flexible manner.

Pinto (1998, in preparation) found that students made definitions operable in (at least) two distinct ways, which she termed:

- *giving* meaning *to the concept definition from concept imagery,*
- extracting meaning *from the concept definition through using it to make formal deductions.*

It transpires that these two routes can *both* be successful and both be meaningful, but that they cause different difficulties at different times in their development. *Giving meaning* can often lead to failure because the student works with the ideas that are *believed* to be true because of previous experience and fails to integrate this meaning into the step-by-step proof. For this approach to work, the student must be constantly reconstructing personal imagery to take into account new formal ideas. It can produce mathematicians of the highest quality, but it can also be a high-risk strategy. *Extracting meaning,* on the other hand, requires the student first to play with the definition, to attempt to see what can be deduced from it, gaining meaning from this activity. Again, this is a difficult strategy. Many students are not able to cope with the meanings of the definitions separate from their concept imagery and can make little progress at all, particularly where several quantifiers are involved.

A case in point is the definition of convergence of a sequence in the form:

Given $\varepsilon>0$, $\exists\ N\in \boldsymbol{N}$ such that $n\geq N$ implies $|a_n-a| < \varepsilon$.

Pinto found a spectrum of different routes taken by students in coping with this concept definition and its use in formal proof. We illustrate these beginning with examples of those who fail.

Giving meaning unsuccessfully

When presented with a sophisticated idea, such as the notion of convergence of a limit, many students simply talk about their concept imagery. Laura, a student teacher, who spends much of her time teaching children by example, evoked many personal images of the limit concept:

> "The number where the sequence gets to, but never quite reaches."
>
> Let a_n be the sequence and L is the limit which it tends to. Then when some initial values are placed into the formula of the sequence the answers will never reach the value of L (negative or positive).
>
> "... oh, yes, I put 'never reach', and it *can* reach, and that will be the limit of it. ..."
>
> "... But it won't never get bigger than the limit. The limit is like the top number it can possibly reach. And I put never reach."
>
> (Laura, selected comments from her first interview)

She was unable to write down the definition in any formal sense, although she had mental pictures that gave her meaning for some of the theorems. She could "see" limiting notions in a dynamic, idiosyncratic manner, but not prove them.

Extracting meaning unsuccessfully

Rolf attempted to learn the definition. He thought he had remembered it properly, but was mistaken:

> "Umm ... I wrote it many times because we use it all the time, every time we are asked a question we have to use and that's how I remembered it. *I don't think I will ever forget it now*. We have done it so many times."

He wrote:

$(a_n) \to L$ if

$|a_n - L| < \varepsilon$ $\quad \varepsilon \in \mathbb{Z}$, $\varepsilon > 0$ $\quad$ for $n \geq N$

(Rolf, first interview)

He was consequently unable to use the definition coherently and, when asked to write a statement saying a_n does *not* tend to the limit *L*, wrote

$|a_n - L| > K$ for all $n \geq N$

$k > 0$

(Rolf, second interview)

Other students often worked from the inner quantifier out, some negating the definition simply by changing round the inequality in the inner statement from "<" to ">"or "≥", as in the following example:

A sequence a_n does not tend to the limit L if for any $\varepsilon > 0$, there exists a positive integer N s.t.

$|a_n - L| > \varepsilon$, whenever $n \geq N$

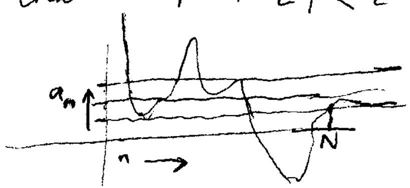
(Robin, second interview)

Others coped with one or more quantifiers, working from the inner quantifier out, as noted by Dubinsky *et al.*, (1988). But not all students did this. For instance, Chris read the definition as it came from left to right and made sense of it by manipulating his concept imagery, as we shall now see.

Examples of successful students giving and extracting meaning

Chris was a remarkable student who built his definitions by giving meaning from his imagery:

~~If $a_n \to L$ then there~~ exists
For all $\varepsilon > 0$, there exists $N \in \mathbb{N}$
such that $|a_n - L| < \varepsilon$ for all $n \geq N$



He said:

> "I don't memorize that [the definition of limit]. I think of this [picture] every time I work it out, and then you just get used to it. I can nearly write that straight down." (Chris, first interview)

As he drew the picture, he gestured with his hands to show that first he imagined how close he required the values to be (either side of the limit), then how far he would need to go along to get all successive values of the sequence inside the required range:

> ”I think of it graphically … you got a graph there and the function there,and I think that it’s got the limit there … and then ε, once like that, and you can draw along and then all the … points after N are inside of those bounds. … When I first thought of this, it was hard to understand, so I thought of it like that’s the n going across there and that’s a_n. … Err, this shouldn’t really be a graph, it should be points.” (Chris, first interview)

Ross, on the other hand, took a formal approach, extracting meaning from the formal definition. He explained that he learns the definition:

> “Just memorizing it, well it’s mostly that we have written it down quite a few times in lectures and then whenever I do a question I try to write down the definition and just by writing it down over and over again it gets imprinted and then I remember it.”
>
> (Ross, first interview)

and wrote:

A sequence (a_n) tends to limit L if, $\forall\ \varepsilon>0,\ \exists\ N(\varepsilon)\in\mathbb{N}$ s.t. $\forall\ n\geq N$;

$$|a_n - L| < \varepsilon.$$

(Ross, first interview)

The difference between *giving* and *extracting* meaning arose again when the students were asked to say what it means for a sequence *not* to converge. Chris was an exceptional student. From his long experience working mentally with the concept, he was able to perform a direct thought experiment and write the definition of non-convergence straight out:

A sequence (a_n) does not tend to a limit if for any L, there exists $\varepsilon > 0$ such that $|a_n - L| \geq \varepsilon$ ~~whenever n ≥ N, for~~ for some $n \geq N$ for all $N \in \mathbb{N}$

(Chris, second interview)

Ross, an extractor of meaning, first wrote the definition of convergence:

$$\forall\, \varepsilon>0,\ \exists\, N(\varepsilon)\in\mathbb{N} \text{ s.t. } \forall\, n\geq N.$$
$$|a_n - L| < \varepsilon.$$

(Ross, second interview)

and then negated successive quantifiers formally:

$$\forall\, L,\ \exists\, \varepsilon>0 \text{ s.t. } \forall\, N(\varepsilon)\in\mathbb{N}\ \exists\, n\geq N, \text{ s.t. }$$
$$|a_n - L| \geq \varepsilon$$

(Ross, second interview)

Note that both students made errors at various times. When explaining convergence, Chris drew a continuous curve instead of discrete points. However, he was not focusing on the precise nature of the drawing, rather on the *behavior* of the sequence as it moved about up and down, eventually getting within a prescribed range of the limit. As part of his "sense-making", he considered whether increasing N caused ε to become smaller, before settling on the idea that he must first specify ε and use this to determine N. He experimented and modified his ideas over several weeks, building a highly connected conceptual structure.

Ross made a subtle error in his negation by writing "for all N(ε)", although N no longer depends on the (fixed) ε. This arose because he had written "∃ N(ε)" in the definition to give it greater meaning than simply "∃ N". Had he written the latter, the error would not have occurred (or would not have been noticed).

Even *successful* students make errors as they attempt to come to terms with complex ideas. They become more successful because they have the will and stamina to overcome these setbacks.

The move to formal mathematics can occur in different ways. Some students build up powerful, flexible imagery, capable of being used to suggest and underlie formal proofs; others attempt to build a consistent theory from the definition in a formal way. Neither of these routes is easy, and many fail.

Summary

Although the experts in mathematics may claim to share a coherent notion of proof, the cognitive development of proof is dependent on the cognitive structure and representations available to the learner at a given time. The formal concept of proof in terms of definition and logical deduction has a significant cognitive difficulty; it requires a reversal from "concepts described verbally" to "verbal definitions which prescribe concepts". This is likely to be highly confusing to non-experts.

The loss of Euclidean geometry in the UK National curriculum has removed any suggestion of a global mathematical theory built from explicit deductive foundations, replacing it mainly by the use of algebra to express generalities.

The cognitive development of students needs to be taken into account so that proofs are presented in forms that are potentially meaningful for them. This requires educators and mathematicians to rethink the nature of mathematical proof and to consider the use of different types of proof related to the cognitive development of the individual.

References

Barnard, A. D. & Tall, D. O., (1997). Cognitive Units, Connections and Mathematical Proof, *Proceedings of P. M. E. 21*, Lahti, Finland.

Bills, E. & Tall, D. O., (1998). Operable Definitions in Advanced Mathematics: The Case of the Least Upper Bound, *Proceedings of P. M. E. 22,* Stellenbosch, South Africa. (to appear).

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

Cornu, B, (1991). 'Limits'. In D. O. Tall (ed.), *Advanced Mathematical Thinking*, (pp.153-166), Kluwer: Dordrecht.

Dubinsky, E., Elterman, F. & Gong, C., (1988). The student's construction of quantification, *For the Learning of Mathematics*, 8, 44–51.

Gray, E. M. & Tall, D. O., (1994). Duality, Ambiguity and Flexibility: A Proceptual View of Simple Arithmetic, *The Journal for Research in Mathematics Education*, 26 (2), 115–141.

Hadamard, J., (1945). *The Psychology of Invention in the Mathematical Field*, Princeton University Press, (Dover edition, New York 1954).

Nunn, T. P., (1914). The Teaching of Algebra (including Trigonometry). Longmans: London.

Poincaré, H., (1913). *The Foundations of Science* (translated by Halsted G. B.), The Science Press, New York (page references as in University Press of America edition, 1982).

Pinto, M. M. F., (1998). *Students' Understanding of Analysis,* University of Warwick, UK.

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

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

Tall, D. O., (1992). The Transition to Advanced Mathematical Thinking: Functions, Limits, Infinity, and Proof, in Grouws D. A. (ed.) *Handbook of Research on Mathematics Teaching and Learning*, Macmillan, New York, 495–511.

Tall, D. O., (1995). Mathematical Growth in Elementary and Advanced Mathematical Thinking. In L. Meira & D. Carraher, (Eds.), *Proceedings of PME* 19, Recife, Brazil, I, 61–75.

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

Van Hiele, P. & D., (1959). *The child's thought and geometry*. Reprinted (1984) in D. Fuys, D. Geddes & R. Tischler (Eds.), *English translation of selected writings of Dina van Hiele-Geldof and Pierre M. van Hiele*, (pp. 1–214). Brooklyn NY: Brooklyn College.

Van Hiele, P., (1986). *Structure and Insight*. Orlando: Academic Press.

Vinner, S. & Hershkowitz, R., (1980). Concept images and some common cognitive paths in the development of some simple geometric concepts, *Proceedings of the Fourth International Conference of P. M. E.*, Berkeley, 177–184.