Cognitive development, representations and proof

David Tall

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

The purpose of this paper is to highlight the different forms of proof afforded by different types of mathematical representation. The form of proof generally accepted by mathematicians is logical proof with formal concept definitions and deductions using the predicate calculus (although there are many subtle differences in acceptability of a proof in the mathematical community). However, the cognitive development of a notion of proof must take into account the differing forms of representation available to the learner at various levels of sophistication. In particular, there are two very different parallel developments of visualisation and symbolisation with different forms of proof.

Introduction

"Proof" is regarded as a central concept in the discipline of mathematics. It is important for two reasons.

(1) (*Local*) Based on explicit hypotheses, a proof shows that certain consequences follow logically,

(2) (*Global*) Such logical consequences themselves can be used as "relay results" (Hadamard 1945) to build up mathematical theories.

In the recent past (eg since the mid nineteenth century in England), Euclidean geometry has been considered as an introduction to both (1) and (2). However, this has fallen out of favour because of the difficulties encountered by children (eg Senk, 1985, 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). The NCTM standards in the USA suggest that there should be increased attention on short sequences of theorems and decreased attention to Euclidean geometry as an axiomatic system, favouring (1) over (2). In England the demise of geometry has proceeded further. The Association for the Improvement of Geometry Teaching (which later became the Mathematical Association) was

This paper was prepared for the conference on *Justifying and Proving in School Mathematics*, Institute of Education, London, December 1995, pp. 27–38.

formed in 1871 to address the problem of making geometric proof meaningful, a task that was never satisfactorily completed. Geometry stayed as an equal partner with arithmetic and geometry in secondary school until the mid twentieth century. It has since been replaced in the English National Curriculum by the study of “Shape and Space” with Euclidean proof only being mentioned in passing. For instance, in the final level of the Curriculum, attainment target 10 of “Shape and Space” mentions

> knowing and using angle and tangent properties of circles,

and “Using and Applying Mathematics” suggests under the heading “handling proof and definition” 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.

Thus it is that geometry as the introduction to a global deductive system in school has been replaced by occasional references to knowledge and use of facts and investigations of how to prove isolated theorems. Attention is now firmly focused in school on the local aspects of proof. Moreover, in investigations in the British curriculum, it is changing its style to expressing generalisations, often in algebraic form.

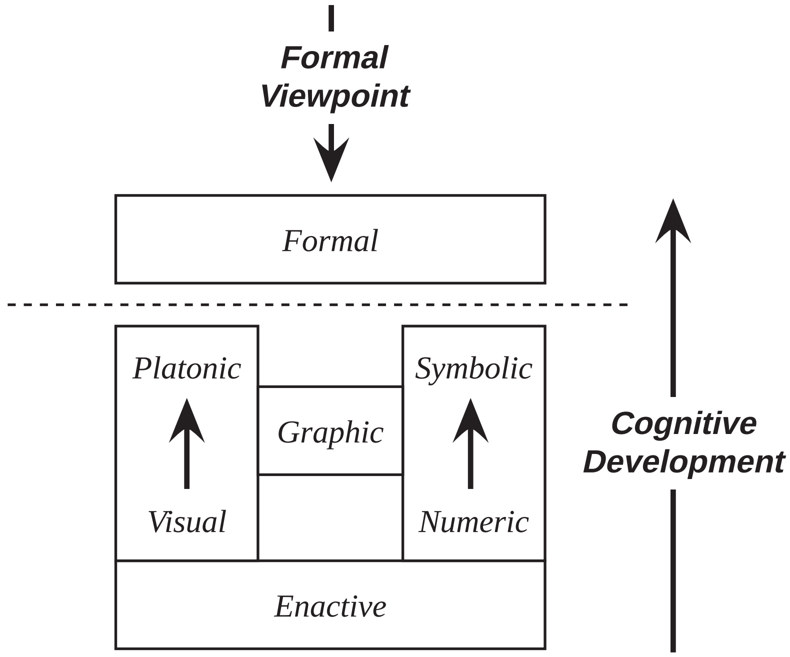
This change in emphasis from verbalising visual ideas in geometry to generalising arithmetic ideas in algebra requires us to consider carefully the meaning of the term “proof” and its development from short deductions in various areas of elementary mathematics to the global structure of formal proof in advanced mathematics.

Even at the formal level, the use of the single word “proof” disguises the fact that there are many different views of proof, dependent on different historical and cultural contexts. For instance, Cantor’s elegant proof that there exists a real number that is not the solution of an algebraic equation was not accepted by many of his contemporaries. It was rejected for publication in *Crelle’s Journal* by Kronecker in 1873, because the proof presented a counting argument that there are “more” real numbers than solutions of algebraic equations but failed to *construct* one.

By acknowledging that different standards and types of proof exist even at a formal level, we can begin to appreciate that different forms of proof are likely to be appropriate in different contexts. In this paper, consideration is focused on different forms of proof which might be appropriate at various stages of cognitive development, dependent on different representations of knowledge that may be available.

Cognitive development of representations and proof

In Tall (1995), following the ideas of Bruner (1966), I outlined how *enactive* representations based on interactions with the environment and communication through action and gesture provide a foundation for mathematical growth, and how *visual* and *symbolic* representations reveal differing kinds of development which interact with each other and, at an advanced level, give rise to the need for *formal* definition and proof. These can be represented in outline by the following diagram:



Cognitive development of representations

At the foundation is enactive interaction with the environment. On the visual side the representations take on successively more subtle meaning. Visual concepts begin as gestalts whose meaning is refined through interaction, verbal description and discussion. In geometry, figures are classified into distinct types (circles, squares, rectangles, etc) then into hierarchies (squares as a subset of rectangles as a subset of quadrilaterals, etc), then developing more formal definitions and deductions in Euclidean geometry (Van Hiele, 1959, 1986). These follow a sequence of stages where the meaning changes, first with objects as perceived physical examples, becoming cognitively more abstract through the use of language and imagination, so that a physical "straight line", which can never be drawn precisely, becomes a perfect mental object having no width, being perfectly straight, and extensible at will in either direction. Thus visual representations initially represent physical objects but through cognitive development they become *platonic mental objects*, the “perfect” abstract counterparts of physical experience.

On the other hand, in arithmetic and algebra the symbols are designed for calculation and manipulation and derive their great power from the fact that they not only evoke a process, such as addition, 2+3, but also a concept, the sum 2+3 is 5. This use of a symbol to evoke either *pro*cess or *con*cept is called a *procept* (Gray & Tall, 1991, 1994). Procepts have power because they evoke processes to *do* mathematics and concepts to *think* about mathematics. In the diagram I have used the terms “numeric” and “symbolic” to stand for the increasing sophistication from the numeric procepts in arithmetic to the symbolic procepts of algebra and the even greater generality of symbolic functions.

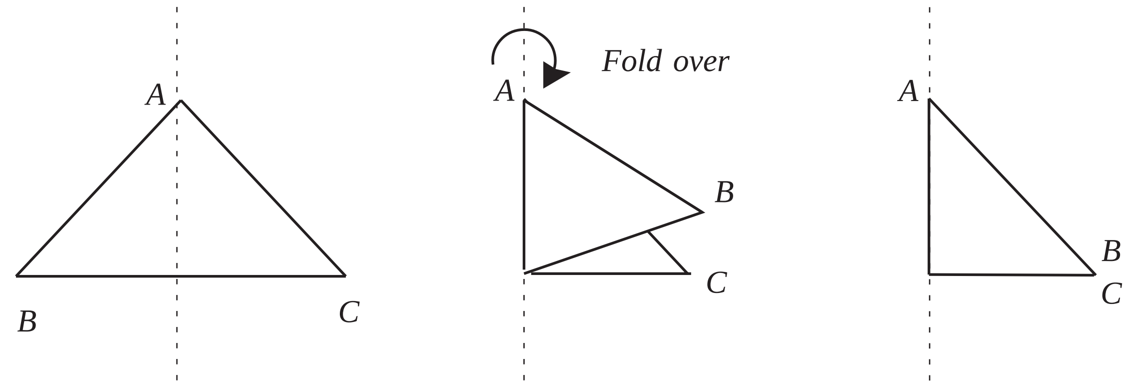
Straddling the visual/platonic and the numeric/symbolic are *graphic* representations, used here in the technical sense of linking together visualisation and symbolisation, for instance, through visual representations of numerical relationships or the use of the real line and the coordinate plane to visualise symbolic relationships.

Separate from these representations are the formal representations of mathematics through definitions and deductions using formal proof. The move from visual and symbolic representations to formal representations requires a huge cognitive reconstruction. In elementary mathematics, concepts are developed and then described verbally and represented visually. The concept *precedes* the description. If there is a mismatch, it is the *description* which is (usually) changed, not the concept. In the formalisation of advanced mathematics, it is the *definition* which comes to have primacy. The formal concept is *constructed from the formal definition*, and the properties of the formal object are *only* those which can be deduced from the definition. The significant reconstruction required to establish the definition as the basis of formal concept construction is signified in the diagram by a dotted line which separates elementary mathematics from the advanced form of thinking required in formal mathematics.

From the viewpoint of the expert, formal proof is meaningful and essential, but this depends on the cognitive growth of the expert. It cannot be understood by those who lack the necessary sophistication. It cannot even be an objective for the learner to reach because it, as yet, has no meaning for the learner. Proof is *context dependent*, and its cognitive development must take account the cognitive structure and representations available to the growing individual.

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, 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.

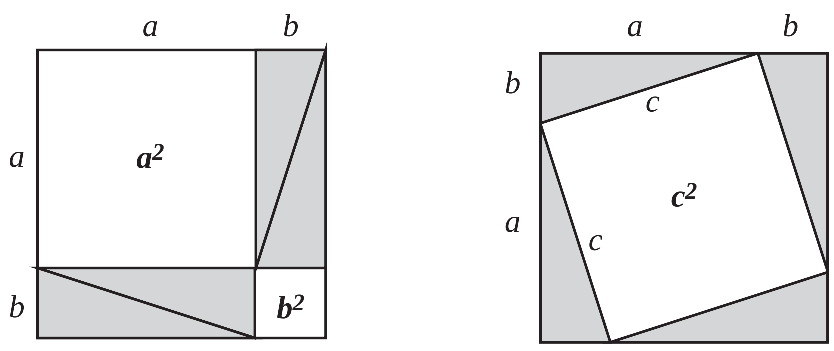


An enactive (visual) proof that equal sides imply equal angles

Such a proof invariably involves either specific examples, or specific examples seen as representative *prototypes* of a class of examples.

Visual Proof

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 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 a single square area c^2, giving $a^2 + b^2 = c^2$.

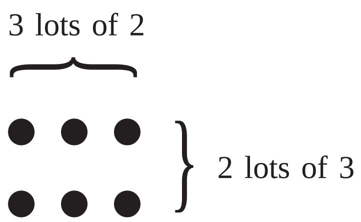


An (enactive) visual proof of Pythagoras (after Bhaskara)

To "see" this proof, it is essential to be able to imagine how the triangles can be moved around from one configuration to another.

Note that any actual drawing will have specific 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".

In the same way many arithmetical statements can be "seen" to be true by using visual configurations in a generic way as prototypes for a class of statements. 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:



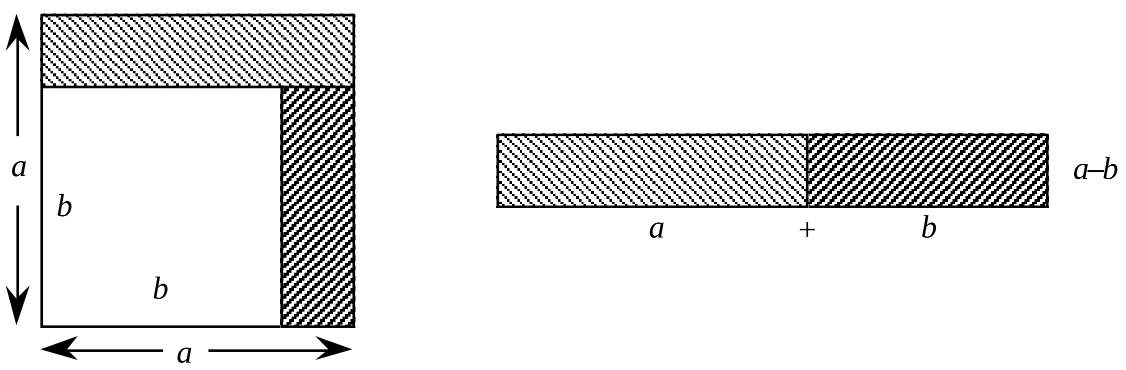
a graphic proof that 3×2 is the same as 2×3

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 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 visualised (at least for positive a and b) in the following generic diagram:



Taking a square side *b* from a square side *a* and rearranging what is left as (a–b)×(a+b)

This graphic proof again has elements of enaction to see the dynamic rearrangement of the parts.

Manipulative proof

The previous algebraic identity can also be proved by manipulation. To show that

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

all that is necessary is to multiply out the brackets on the left hand side and cancel the terms *ba* and *–ab*.

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.

However, *generic* proof is possible, where a specific statement is seen to be typical of a class of statements. For instance, to show that the square of any number cannot be 2, one might note that any fraction in lowest terms can be factorised 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 factorisation 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 factorises as 2/1 and has an *odd* number of 2’s in the numerator.

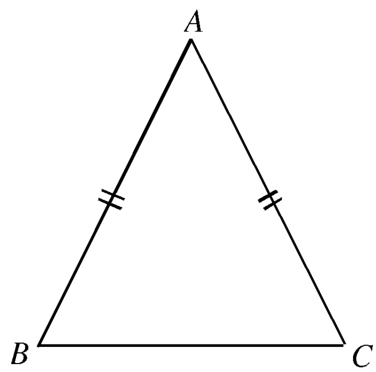
In Tall (1979) I showed that students in the first year of university expressed a strong preference for the generic proof over the standard proof by contradiction, but note that this does *not* mean that the generic proof is preferable in the long term. Proof by contradiction is an essential element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

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).

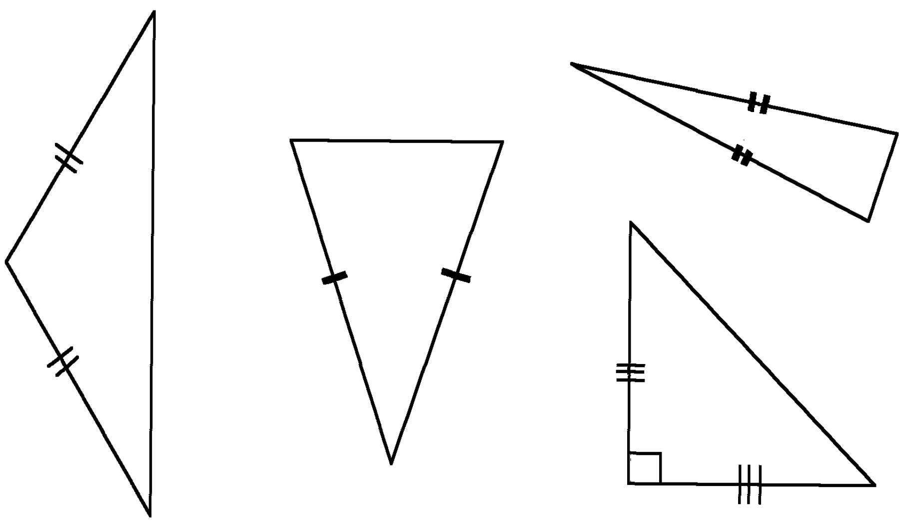
This is the most commonly occurring method of "proof" in the English National Curriculum, and occurs widely in numerical investigations. However, it involves meaningful manipulative facility in algebra rather than logical deduction. In general logic has a lower (almost non-existent) priority in the National Curriculum compared with the curriculum of other countries (eg Italy).

Euclidean Proof as a verbal translation of generic visual proof

The ideas of Euclidean geometry are inspired by visual representations but they are formulated verbally to give the proofs greater generality. A theorem in Euclidean geometry specifies a certain geometric configuration. A figure drawn to accompany the theorem is a generic picture which represents *any* configuration satisfying the statement. The verbal proof then 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 if ΔABC has $AB=AC$ then $\angle B=\angle C$ applies not just to this triangle ABC:



but to all these triangles too:



In this way Euclidean proof is verbal generic proof applying to the whole class of geometric figures having the given properties.

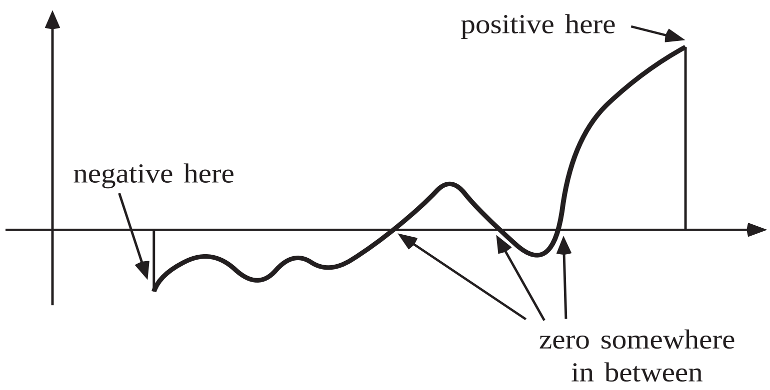
Weaknesses in Visual Proof

In the nineteenth century it was realised 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 may prove to be unsatisfactory later on.

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*.



The intermediate value theorem

Enactively the notion of "continuous function" is something that is drawn "continuously" without take the pencil from the paper. Physically, 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

When formal definitions and deductions are introduced, the intermediate value theorem proves to be *true* when the graph is defined on an interval of the real numbers, but *false* when applied to a function defined only on the rational numbers. (The function $f(x)=x^2-2$ is negative for $x=1$, positive for $x=2$, but not zero for any *rational* number in between.) Since visually we cannot distinguish between the real and the rational number line, it becomes clear (to an expert) that something more is required for a formal proof, namely, a logical sequence of deductions described verbally starting from verbal definitions.

The change to the formal level requires a huge cognitive struggle. First is the difficult reversal of primacy of definition and concept, so that the concept definition is used to *define* the concept, not just to describe some of its salient features. Then there is the substantial problem of identifying the concepts in one's cognitive structure that depend only on logical deductions from definitions and only to use these links as stepping stones in mathematical proof.

Difficulties occur when the enactive or visual form of the proof does not suggest an obvious sequence of deductions to use for a formal proof, so that the individual seems to "know" that the theorem is true and yet has no method of *proving* it. There are numerous examples in topology where an "obvious" visual property fails to have a correspondingly simple proof (such as 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 the outside).

Exacerbating the situation is the often great complexity of quantifiers which occur 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. The overburdening detail of a strict formal proof is explicitly or implicitly suppressed to give socially acceptable forms of proof in the mathematical community.

For the learner there are other stages of difficulty in formal proof, for instance, the use of proof by contradiction to prove an existence theorem which fails to *construct* the mathematical object which hypothetically exists. Or the potential infinity of steps in an elementary induction proof, compared with the finite use of the induction axiom in the Peano postulates.

Summary

Although the experts in mathematics may claim to share a more or less coherent complex of ideas about the nature 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 requires a cognitive reversal from "concepts described verbally" to "verbal definitions which prescribe concepts" which is unlikely to be fully available to less experienced individuals. The formal concept of proof is therefore likely to be highly confusing to non-experts (and, I would suggest, has a far less clear corporate meaning than the mathematical community might care to claim).

The loss of Euclidean geometry in the UK National curriculum has removed any suggestion of a global mathematical theory built from explicit deductive foundations. This has largely been replaced by a manipulative form of algebraic proof which lacks the logic of true deduction.

Educators and mathematicians need to rethink the nature of mathematical proof and give appropriate consideration to the different types of proof related to the cognitive development of the individual.

References

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

Gray, E. M., & Tall, D. O. (1991). Duality, ambiguity and flexibility in successful mathematical thinking. In F. Furinghetti (Ed.), *Proceedings of the Fifteenth International Conference for the Psychology of Mathematics Education,* (pp. 72–79). Assisi, Italy.

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).

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

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.