De Villiers, M. The Role of Proof in Investigative, Computer-based Geometry: Some personal reflections. Chapter in Schattschneider, D. & King, J. (1997). *Geometry Turned On!* Washington: MAA, pp. 15-24.

The role and function of proof in dynamic geometry: Some personal reflections

Michael de Villiers, Mathematics Education,
University of Durban-Westville, South Africa

Introduction

There is a tendency amongst many mathematicians to report only their final results in a neatly organized fashion, not discussing and reflecting much upon the processes of discovery/invention and proof. This tends to give a distorted perspective of mathematical creativity as being always purely deductive.

In what follows, I intend to heuristically present some personal discoveries I recently made with the aid of dynamic geometry software. As far as I know these results are original and have not been published elsewhere before. In all these examples, actual conviction preceded the eventual proofs, and the purpose is to reflect on the need and function of proof in such cases. Readers are also encouraged to first investigate the results with their own dynamic constructions before reading the proofs. A personal model of how new discoveries are sometimes made in mathematics, as well as its relevance to mathematics education, will also be briefly discussed.

Example 1

The false impression is sometimes created that mathematicians are only problem solvers who spend most of their time trying to solve *already given* problems. However, mathematicians continually create their own new problems by asking questions, making hypotheses and testing them.

A problem-posing heuristic I have often found useful is to ask "*what - if?*" questions when coming across any mathematical result. Consider for example Figure 1 which shows a convex quadrilateral ABCD with equilateral triangles ABP, BCQ, CDR and DAS constructed on the sides so that the first and third are exterior to the quadrilateral, while the second and the fourth are on the same side of sides BC and DA as is the quadrilateral itself. Then quadrilateral PQRS is a parallelogram.

When I first came across this result in Yaglom (1962:39), I immediately wondered what *would happen if* instead of equilateral triangles, *similar* triangles PBA, QBC, RDC and SDA were constructed on the sides. After quickly constructing a dynamic version of the figure with *Cabri* by using a macro-construction for the similar triangles on the sides as shown in Figure 2, it was found that PQRS remains a parallelogram. By measuring opposite sides and dragging the vertices of the figure around the screen, it can be seen that this result appears to be true in general. In fact, I used the property checker of *Cabri* to check whether both pairs of opposite sides were parallel, and each time it gave the message: "*this property is* ***true in a general position***"[1].

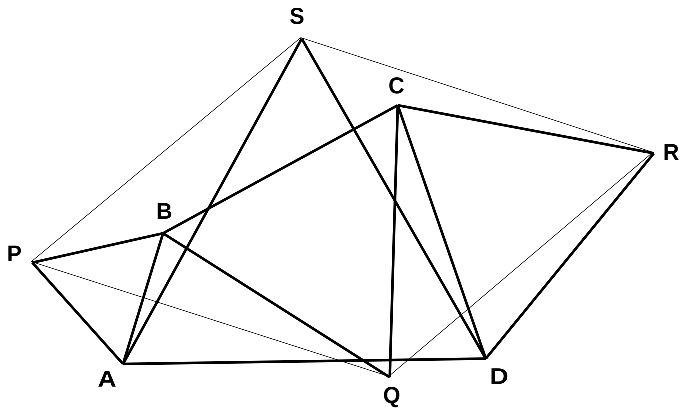

Figure 1

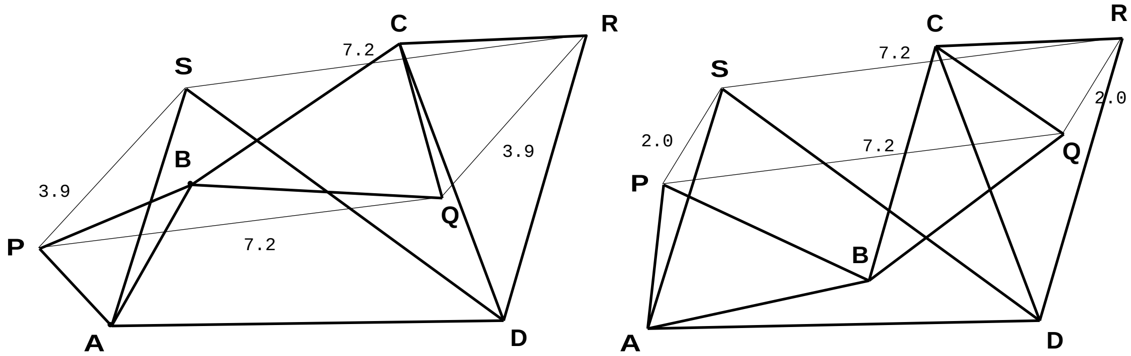

Figure 2

Although I was initially very skeptical about the property checker on *Cabri*, I have yet to "*catch*" it out and have consequently learnt to trust it to a very large degree when investigating new or unknown conjectures like that above. Armed with conviction that the generalization was indeed true[2], I then proceeded with the task of constructing a deductive proof. Why did I still feel a need to prove the above result if I was already convinced of its truth?

Firstly, it is important to point out that it is precisely because I was convinced of its truth that I felt challenged to find a deductive proof, not because I doubted the result. Why? Well, here was a result that was obviously true and I was intrigued to try and find out *why* it was true. If the above construction had not shown PQRS to be a parallelogram, I would certainly not have wasted my time trying to find a proof. I therefore experienced the search for and eventual construction of a deductive proof in this case as an intellectual challenge, definitely not as an epistemological exercise in trying to establish its "*truth*".

Proof

The proof below uses the concept of spiral similarity. A *spiral similarity (k,x)* is an *enlargement* or *reduction* from a point with a factor *k*, followed by a *rotation* around the

same point through the angle *x*, or vice versa. The proof also uses the following Lemma which is proved in De Villiers (In press) or is left as an exercise for the reader:
"The sum of two spiral similarities (k,α) and $(1/k,\beta)$ about centres O_1 and O_2 is equivalent to a translation if α and β are equal in size and opposite in direction."

Consider Figure 3. The sum of the four spiral similarities (*k;z*), (*1/k;z*), (*k;z*) and (*1/k;z*) around centres M_1, M_2, M_3 and M_4, where the direction of the first and third rotations is opposite to that of the second and fourth, carries the vertex A of the quadrilateral onto itself. (Note that $k = M_1B / M_1A = M_2B / M_2C = M_3D / M_3C = M_4D / M_4A$ from the similarity of the triangles). But according to the Lemma above, the sum of the two spiral similarities about M_1 and M_2 is a translation of M_1 by the segment $M_1M_1^{'}$ where $M_1^{'}$ is a vertex of the triangle $M_2M_1^{'}M_1$, which is similar to the triangles M_1AB, M_2CB, etc. constructed on the sides of the quadrilateral $(M_2M_1 / M_2M_1^{'} = k, \angle M_1M_2M_1^{'} = z$ and the direction of rotation from M_2M_1 to $M_2M_1^{'}$ coincides with the direction of rotation from M_2B to M_2C).

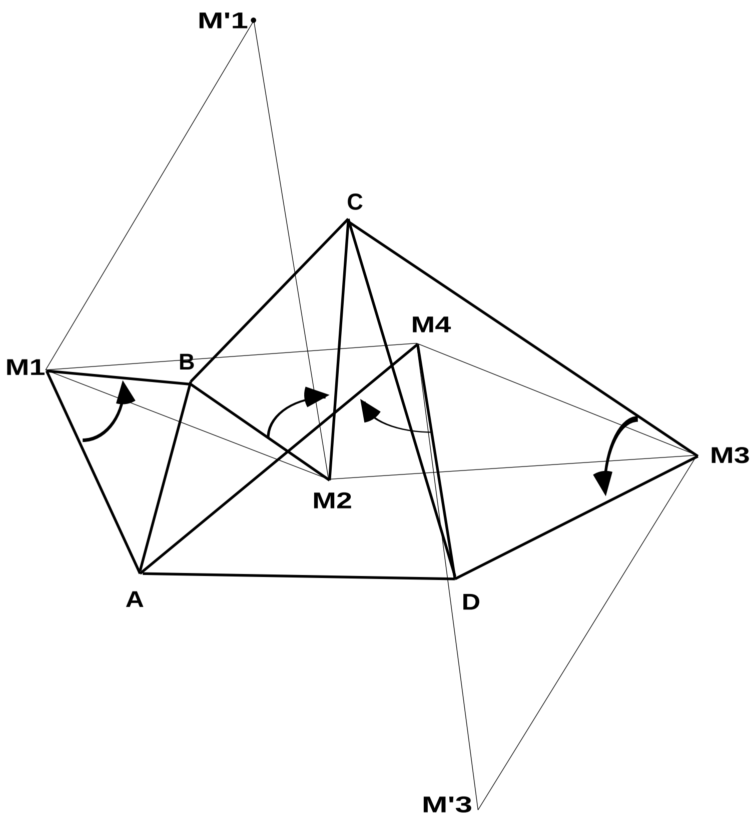

Figure 3

Similarly the sum of the two spiral similarities about M_3 and M_4 is a translation given by the segment $M_3M_3^{'}$ where triangle $M_4M_3^{'}M_3$ is similar to triangles M_1AB, M_2CB, etc. and also to $M_2M_1^{'}M_1$ (the direction of rotation from M_4M_3 to $M_4M_3^{'}$ is the same as the direction of rotation from M_4D to M_4A).

Thus the sum of the two translations given by the segments $M_1M_1^{'}$ and $M_3M_3^{'}$ carries the point A into itself. But if the sum of two translations leave even one point fixed, then this sum must be the identity transformation, that is, the two segments that determine the two translations must be equal, parallel and oppositely directed. But if $M_1M_1' = M_3M_3'$ then the similar triangles $M_2M_1'M_1$ and $M_4M_3'M_3$ are congruent, and therefore $M_1M_2 = M_3M_4$. Since the two congruent triangles $M_2M_1'M_1$ and $M_4M_3'M_3$ are furthermore so situated that M_1M_1' is parallel and oppositely directed to M_3M_3' it follows that M_1M_2 and M_3M_4 must also be parallel and oppositely directed, and therefore that the quadrilateral $M_1M_2M_3M_4$ is a parallelogram.

Example 2

Some years ago I came across Von Aubel's theorem in Martin Gardner's book **Mathematical Circus** (1981:176-179), namely, that the centres of squares on the sides of any quadrilateral ABCD, form a quadrilateral EFGH with equal and perpendicular diagonals (see Figure 4). Again I wondered what would happen if instead of squares on the sides, one constructed similar rectangles or rhombi on the sides. It was however not until recently that I had an opportunity to investigate these questions.

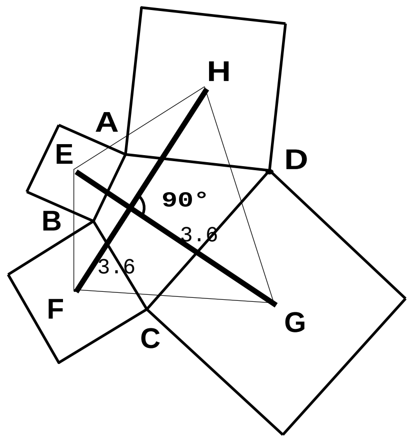

Figure 4

After some initial experimentation on the arrangement of the similar rectangles and rhombi on the sides, the following two generalizations of Von Aubel were discovered:

1. If similar *rectangles* are constructed on the sides of any quadrilateral as shown in Figure 5, then the centres of these rectangles form a quadrilateral with *perpendicular* diagonals
2. If similar *rhombi* are constructed on the sides of any quadrilateral as shown in Figure 6, then the centres of these rhombi form a quadrilateral with *equal* diagonals.

Again in both cases, it was very easy to click and drag any of the vertices of ABCD around the screen to see if EG remains perpendicular to HF in the first case, and in the second case whether they always remain equal. In fact, I again used the property checker of *Cabri* to verify that both results were indeed "*true in general*"[3].

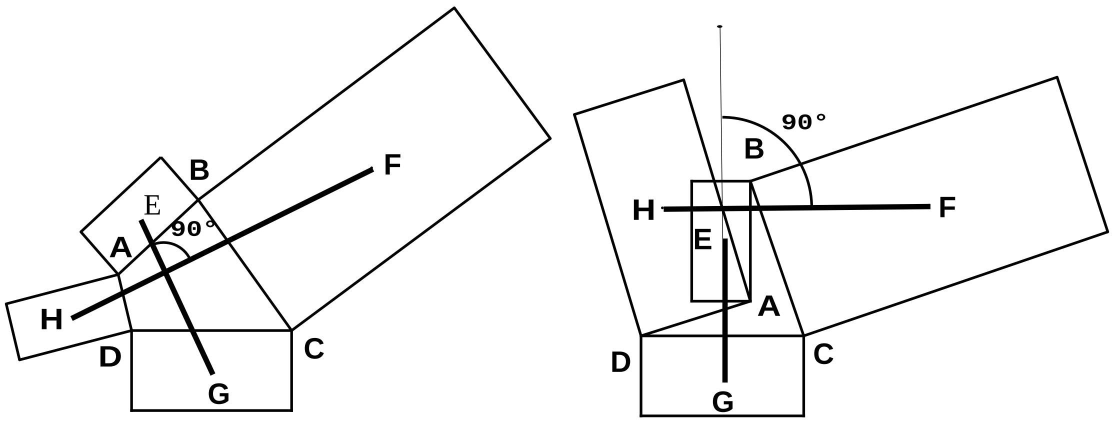

Figure 5

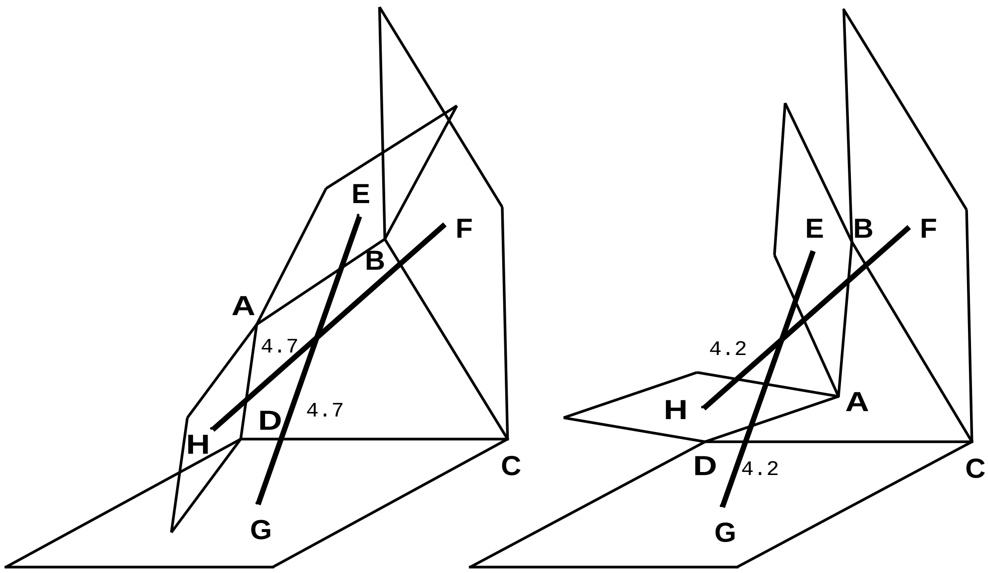

Figure 6

As before, this experimental confirmation motivated me to start looking for proofs. Indeed, since they were clearly true, I reasoned, I should be able to arrive at deductive proofs. In fact, it was precisely this kind of reasoning that kept me motivated after some initial unsuccessful attempts. In other words, it was not a lack of conviction that motivated the continued search for proofs, but in fact its presence. As was the case previously, I therefore did not really experience a need for further certainty, but rather of *explanation* (why were they true?) and of *intellectual challenge* (can I prove them?).

Interestingly, when the base quadrilateral is a parallelogram, the formed quadrilateral in the first case is a rhombus and in the second case, a rectangle. Also note that Von Aubel's theorem is the intersection of the above two results, just as the squares are the intersection between the rectangles and rhombi.

Proofs

The two proofs below also utilize the concept of spiral similarity, as well as the following two Lemmas which are proved in De Villiers (In press) or are left as exercises to the reader:

Lemma 1

"The sum of two rotations with centres O_1 and O_2 through angles α and β respectively is a rotation through the angle $\alpha + \beta$ around a centre O."

Lemma 2

"The sum of two spiral similarities $(k, 90°)$ and $(1/k, 90°)$ around centres O_1 and O_2 is a half-turn around a centre O, the vertex of an isosceles triangle O_1O_2O with $OO_2 = OO_1$."

First result

Consider Figure 7. Let $\angle AM_1B = 2x$, then the sum of the the four rotations about the points M_1, M_2, M_3 and M_4, respectively through angles of $2x$, $180^\circ - 2x$, $2x$ and $180^\circ - 2x$, clearly carries the vertex A of the quadrilateral into itself. It follows that this sum of four rotations is the identity transformation. But the sum of the rotations about M_1 and M_2 is equivalent to a half-turn about O_1 - the vertex of a right triangle $O_1M_1M_2$ with $\angle M_1O_1M_2 = 90^\circ$, since $\angle O_1M_1M_2 = x$ and $\angle O_1M_2M_1 = 90^\circ - x$ (see Lemma 1).

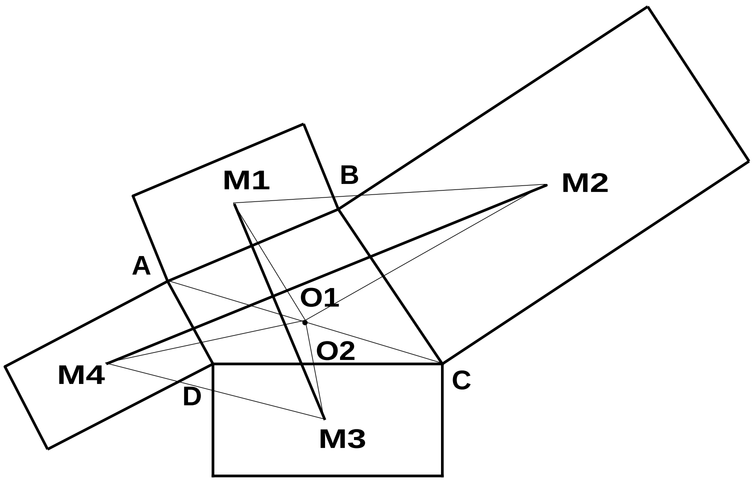

Figure 7

Similarly the sum of rotations about M_3 and M_4 is a half-turn about the vertex O_2 of a right triangle $O_2M_3M_4$ with $\angle M_3O_2M_4 = 90^\circ$. From the fact that the sum of the half-turns about O_1 and O_2 is the identity transformation it clearly follows that these two points coincide.

Since triangles $O_1M_1M_2$ and $O_2M_3M_4$ are similar, we can obtain triangle $O_1M_1M_3$ from triangle $O_1M_2M_4$ by the spiral similarity $(k, 90^\circ)$ about the point $O_1 = O_2$ (a rotation through 90° about the point $O_1 = O_2$ followed by a dilation (enlargement or reduction) from the same point with factor $k = O_1M_1 / O_1M_2 = O_2M_3 / O_2M_4$). Therefore the corresponding segments M_1M_3 and M_2M_4 of triangles $O_1M_1M_3$ and $O_1M_2M_4$ are perpendicular.

Second result

Consider Figure 8. The sum of the four spiral similarities $(k, 90^\circ)$, $(1/k, 90^\circ)$, $(k, 90^\circ)$ and $(1/k, 90^\circ)$ about the points M_1, M_2, M_3 and M_4 clearly carries the vertex A of the quadrilateral onto itself. (From the similarity of the rhombi we have $k = M_1B / M_1A = M_2B / M_2C = M_3D / M_3C = M_4D / M_4A$). It follows that this sum of four spiral similarities is the identity transformation. But the sum of the spiral similarities about M_1 and M_2 is equivalent to a half-turn about O_1 - the vertex of an isosceles triangle $O_1M_1M_2$ with $\angle O_1M_1M_2 = \angle O_1M_2M_1 = \arctan(1/k)$ and $\angle M_1O_1M_2 = 180^\circ - 2\arctan(1/k)$ (See Lemma 2).

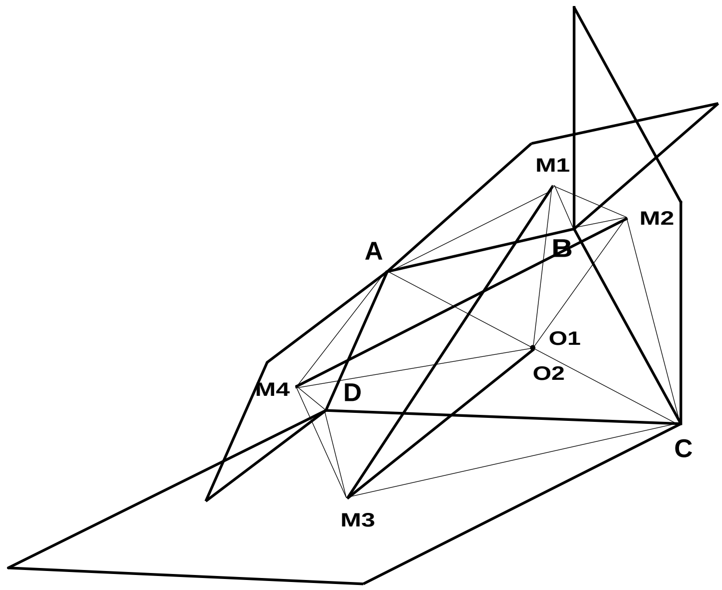

Figure 8

In the same way the sum of the two spiral similarities about M_3 and M_4 is a half-turn about the vertex O_2 of an isosceles triangle $O_2M_3M_4$. From the fact that the sum of the half-turns about O_1 and O_2 is the identity transformation it follows that these two points coincide.
Since triangles $O_1M_1M_2$ and $O_2M_3M_4$ are similar, we can obtain triangle $O_1M_1M_3$ from triangle $O_1M_2M_4$ by a rotation through $180^\circ - 2\arctan(1/k)$ about the point $O_1 = O_2$. Therefore the corresponding segments M_1M_3 and M_2M_4 of triangles $O_1M_1M_3$ and $O_1M_2M_4$ are equal.

It should be pointed out that the above two results can more easily be proved by using vectors (eg. see De Villiers, In press), but it is less explanatory (to me personally) than those given above.

Example 3

In Ross Honsberger's excellent book **Mathematical Gems III** the reader is introduced to the so-called "*equilic quadrilateral*", namely a quadrilateral ABCD with one pair of opposite sides equal, say AD = BC, which are inclined at 60˚ to each other. (The latter condition might also be stated in the form $\angle A + \angle B = 120°$). Then one of the engaging results which is proved is the following:

"If ABCD is an equilic quadrilateral and equilateral triangles are drawn on AC, DC and DB, away from AB, then the three new vertices, P, Q and R are collinear" (see Figure 9).

As before, I again wondered what would happen if ABCD was any quadrilateral with opposite sides equal and triangles PAC, QDC and RDB similar to each other. Would P, Q and R then still be collinear?

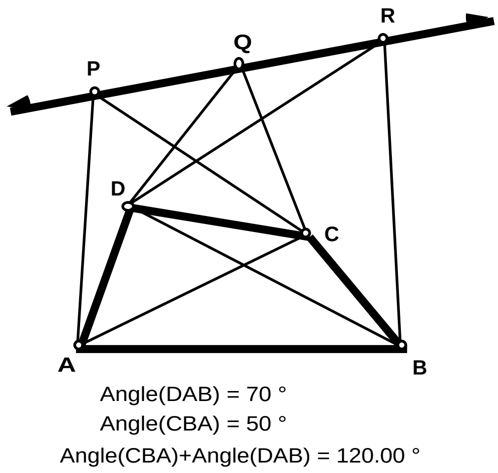

Figure 9

By investigating these questions with *Sketchpad*, I managed to discover the following interesting generalization:

"If similar triangles PAC, QDC and RDB are constructed on AC, DC and DB of any quadrilateral ABCD with AD = BC so that $\angle APC = \angle ASB$, where S is the intersection of AD and BC extended, then P, Q and R are collinear" (see Figure 10).

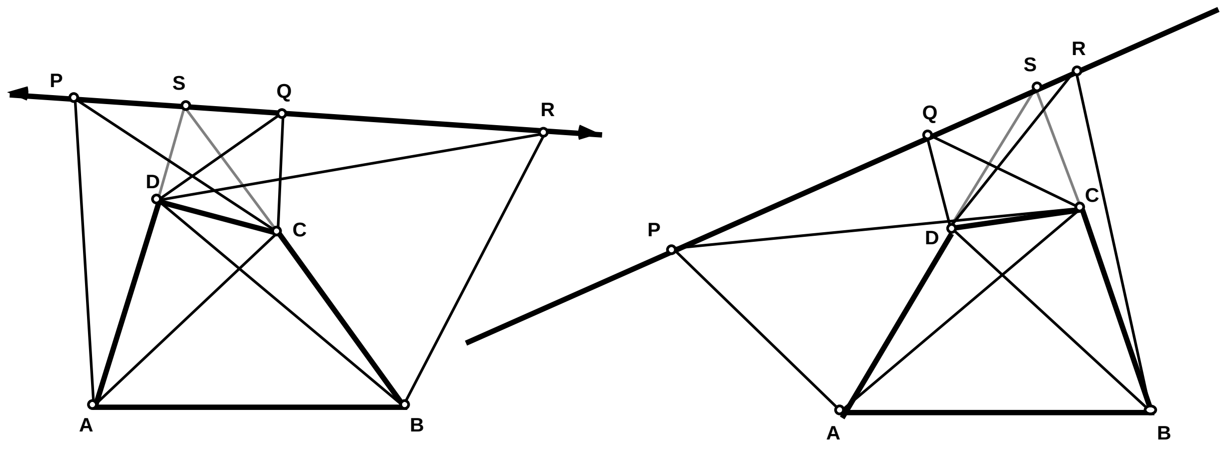

Figure 10

Furthermore, it was found that the point S is also collinear with the other three points. Using a dynamic construction with *Sketchpad* as shown in Figure 10, and varying either angle A or B, or the shape of the similar triangles, it was easy to see that the result was true in general. As before, this conviction was based on the continual experimental confirmation of the result, as well as the absence of counter-examples, and provided the motivation to start looking for a deductive proof; an activity I would hardly have embarked upon had I doubted the result. (In such an event, I would have been looking for counter-examples rather than a proof).

Interestingly, after one or two unsuccessful attempts at proving this result, I then noticed while manipulating the configuration, that $\angle ACB = \angle APQ$, thereby enabling me to construct a proof. This shows how investigation by dynamic software can also sometimes assist in the eventual construction of a proof. Basically the following proof involves first showing $\angle ACB = \angle APQ$ and $\angle ADB = \angle QRB$, and then that PQ and QR have the same direction. It should be noted that the condition that $\angle APC = \angle ASB$ may also be alternatively stated as $\angle PAC + \angle PCA = \angle A + \angle B$ or $\angle APC = 180° - \angle A - \angle B$.

Proof

Consider Figure 11. Connect P with Q and Q with R. Construct CE //= DA as shown. Call the point F where CE cuts AB.

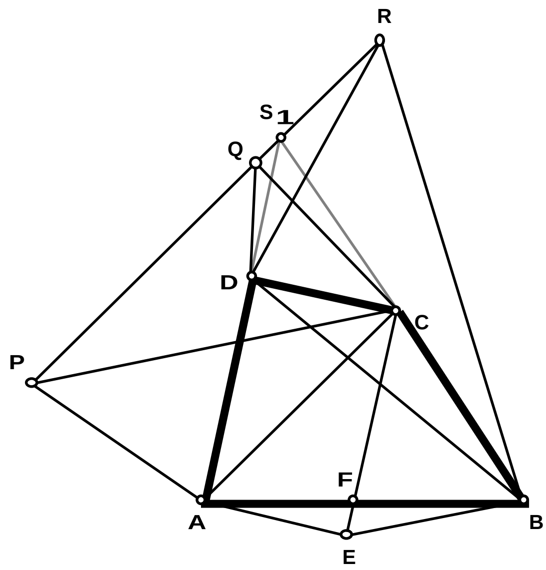

Figure 11

Therefore ADCE is a parallelogram and $\angle CAE = \angle DCA$ (alternate) and $\angle CFB = \angle A$ (corresponding). In triangle CFB we therefore have $\angle ECB = 180° - \angle A - \angle B = \angle APC$.

From the similarity of triangles PAC and QDC, we have $\angle PCA = \angle QCD$. Therefore $\angle PCA + \angle PCD = \angle QCD + \angle PCD$ which implies that $\angle DCA = \angle QCP$, and therefore $\angle CAE = \angle QCP$ ... (1). From the similarity between triangles PAC and QDC we also have:

$$\frac{PC}{AC} = \frac{QC}{DC}.$$

But DC = AE since ADCE is a parallelogram, and therefore:

$$\frac{PC}{AC} = \frac{QC}{AE} \quad ... (2).$$

Thus, according to (1) and (2) above, triangles AEC and CQP are similar. Therefore $\angle ACE = \angle CPQ$. Since $\angle APC = \angle ECB$ as already shown, we have $\angle ACE + \angle ECB = \angle APC + \angle CPQ$ and therefore $\angle ACB = \angle APQ$. By constructing DG //= CB we can similarly prove that $\angle ADB = \angle QRB$.

A counterclockwise rotation of size $\angle PAC$ around A carries C into AP, say C', and B to B'. But since $\angle ACB = \angle APQ$ we have $\angle AC'B' = \angle APQ$ and therefore C'B'//PQ. Thus PQ is inclined to BC at an angle of size $\angle PAC$. Similarly, from a clockwise rotation of size $\angle DBR$ around B, we have that A'D'//QR, that is, QR is inclined to AD at an angle of $-\angle DBR$.

Since AD and BC are inclined towards each other at $180° - \angle A - \angle B = 180° - \angle PAC - \angle DBR$, rotating BC through $\angle PAC$ and AD through $-\angle DBR$ (in appropriate directions), aligns B'C' and A'D' up in the same direction, and we therefore have that PQ and QR also line up in the same direction, i.e. P, Q and R are collinear.

Further construct $\angle QS_1D = \angle QCD$ with S_1 on QR. Connect S_1 with C. We shall now prove $\angle DS_1C = 180° - \angle A - \angle B$ and that S_1CB and S_1DA are straight lines, and therefore that S_1 and S are the same point.

From the construction we have QDCS_1 a cyclic quadrilateral (equal angles on same chord). Thus $\angle DS_1C = \angle DQC = 180° - \angle A - \angle B$ on chord DC. Also $\angle S_1CD = \angle PQD$.

In triangle PQC, we have $\angle PQD = 180° - \angle CPQ - \angle QCP - \angle DQC = \angle S_1CD$. But $\angle BCD = \angle DCA + \angle ACE + \angle ECB$ and $\angle DCA = \angle QCP$ (proved earlier), $\angle ACE = \angle CPQ$ (proved earlier) and $\angle ECB = \angle DQC$ (construction). Therefore $\angle S_1CD + \angle BCD = 180°$ and S_1CB is a straight line. Similarly we can prove that S_1DA is a straight line. Therefore S_1 and S are the same point, namely the intersection of AD and BC. (Note: if S falls on QP we simply construct $\angle QS_1C = \angle QDC$ and show in the same manner that S_1 and S are the same point).

Looking back

Carefully looking back at the above proof, I realized that I had never used the property that AD = BC; in other words, that the result was immediately generalizable to *any* quadrilateral! This illustrates the indispensable value of an explanatory proof which enables one to generalize a result by the identification of the fundamental properties upon which it depends (eg. compare De Villiers, 1990).

The psychology of mathematical discovery and proof

What follows here is a personal model of how new discoveries may sometimes be made in mathematics and is based mainly on the kind of explorations I have done in elementary plane geometry over the past ten years or so, some of which have been described in this paper.

There is no intention however to present it as a model which encompasses the complex totality and rich diversity of mathematical discovery and proof.

Logically, mathematics is assumed to be based upon the following fundamental axiom: *"Something is true (T), if and only if, it can be (deductively) proved (P)"* . However, from a psychological perspective, it is more useful to represent it in the following equivalent, but different logical forms:

(a) the forward implication (T $\Rightarrow$ P): if something is true, then it can be proved.

(b) the converse (P $\Rightarrow$ T): if something has been proved, then it is true.

(c) the inverse (T' $\Rightarrow$ P'): if something is false, then it cannot be proved.

(d) the contrapositive (P' $\Rightarrow$ T'): if something cannot be proved, then it is false.

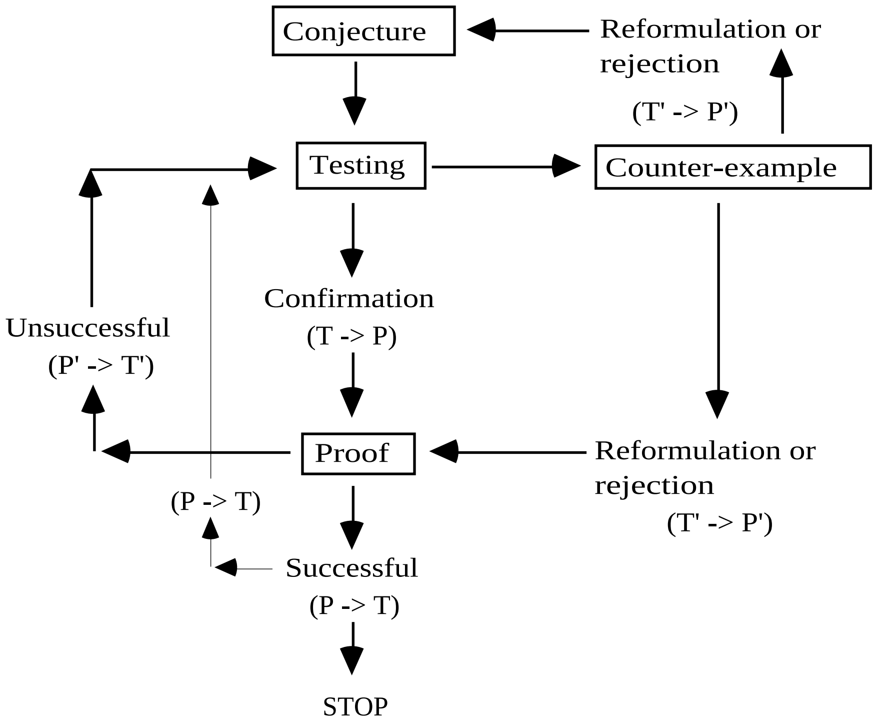

Figure 12

Unfortunately in textbooks and teaching only the converse (P $\Rightarrow$ T) is usually conveyed; in other words, that we must first prove results, before we can accept them as true. However, in actual mathematical research as demonstrated in this paper, the forward implication (T $\Rightarrow$ P), its inverse (T' $\Rightarrow$ P') and contrapositive (P' $\Rightarrow$ T') often play a far greater role in motivating and guiding our actions. For example, suppose we were to make a conjecture and then test it by some cases. If the conjecture is not supported by these cases, we reject it as false and according to the inverse do not even bother trying to prove it. On the other hand, if it is supported by these cases, we might start believing it to be true, which according to the forward implication then gives us the encouragement to start looking for a proof. However, if after a while we are not successful in producing a proof, we might start doubting the validity of the conjecture according to the contrapositive, and then consider some more cases, after which the whole process is of course repeated.

This process of conjecturing, testing, refuting, proving and reformulating can sometimes go through several cycles and is represented in Figure 12. Two famous, often quoted historical examples which spanned many decades and went through several cycles are the Euler-Descartes theorem and Cauchy's theorem about the continuity of the limit of any convergent series of continuous functions.

In the above model, conviction is therefore not seen as the exclusive prerogative of proof nor that the only function of proof is that of verification/conviction. Contrary to popular belief, I do not see proof as an exclusive prerequisite for conviction - to the contrary, as shown in the introductory examples, conviction often precedes proof and is probably far more frequently a prerequisite for the finding of a proof.

One simply does not think: "*Hmm... this result looks very doubtful and suspicious, therefore let's try and prove it*". That would certainly be terribly illogical. For what other, weird and obscure reasons, would we then sometimes spend months or years to try and prove certain conjectures, if we weren't already reasonably convinced of their truth? Sure, to some people these ideas may sound like heresy. I am however not a lone voice crying in the wilderness.

The following quotes by some well-known mathematicians clearly also support the ideas outlined above:

"*...having verified the theorem in several particular cases, we gathered strong inductive evidence for it. The inductive phase overcame our initial* ***suspicion*** *and gave us a strong* ***confidence*** *in the theorem. Without such* ***confidence*** *we would have scarcely found the courage to undertake the proof which did not look at all a routine job. When you have satisfied yourself that the theorem is* ***true****, you start* ***proving*** *it.*" (bold added) - George Polya (1954:83-84)

"*The mathematician at work makes vague guesses, vizualizes broad generalizations, and jumps to unwarranted conclusions. He arranges and rearranges his ideas and becomes* ***convinced*** *of their* ***truth*** *long before he can write down a logical proof. The conviction is not likely to come early - it usually comes after many attempts, many failures, many discouragements, many false starts ...* ***experimental*** *work is needed ... thought -experiments. When a mathematician wants to prove a theorem about an infinite-dimensional Hilbert space, he examines its finite dimensional analogue, he looks in detail at the 2- and 3-dimensional cases, he often tries out a particular numerical case, and he hopes that he will gain thereby an insight that pure definition juggling has not yielded.*" (bold added) - Paul Halmos (1984:23)

"*Actually the mathematician does not rely upon rigorous proof to the extent that is normally supposed. His creations have a meaning for him that precedes any formalization, and this meaning gives the creations an existence or reality ipso facto.*

> *... Great mathematicians know before a logical proof is ever constructed that a theorem must be* ***true*** *...*" (bold added) - Morris Kline (1982:313-314)

Although public conviction through the medium of mathematical journals usually relies heavily on rigorous proof, personal conviction usually depends on a **combination** of intuition, quasi-empirical verification and the existence of some form of logical (but not necessarily rigorous) proof. In fact, a very high level of conviction may sometimes be reached even in the absence of a proof. For instance, in their discussion of the *"heuristic evidence"* in support of the still unproved twin prime pair theorem and the famous Riemann Hypothesis, Davis & Hersh (1983:369) conclude that this evidence is *"so strong that it carries conviction even without rigorous proof."*

Recent years have seen an explosion in the use of computers in many areas of mathematics. For example, Pollak (1984:12) describes the increased use of computers as experimental aids as follows:

> "*...we find ourselves examining on the machine a collection of special cases which is too large for humans to handle by conventional means. The computer is encouraging us to practice unashamedly and in broad daylight, certain customs in which we indulge only in the privacy of our offices, and which we never admitted to students: experimentation. To a degree which never appears in the courses we teach, mathematics is an experimental science...The computer has become the main vehicle for the experimental side of mathematics*".

An interesting development has been the design of a computer programme called *Graffiti* which generates its own hypotheses in graph theory and tests them (see Kolata, 1989). According to the designer of the programme, Fatjlowicz from Houston University, *Graffiti* has already lead to five published articles and at least twenty mathematicians are presently working on proving several other *Graffiti*-results.

Another striking example of the dramatic influence of computer experiments on research in pure mathematics is the sudden burst of developments in areas such as fractal geometry, dynamical systems and chaos. Due to the complexity of the problems in these domains, it was only really with the advent of the computer that mathematicians could start exploring sufficient number of examples to develop their intuition and to see patterns that could lead to conjectures and ultimately to theorems.

Apart from computers being used for experimentation and exploration, they are also increasingly being used as a means of verification. A famous example is of course the computer proof of the four-colour problem in 1976 by Appel & Haken. Recently Branko Grunbaum (1993) used *Mathematica* to verify some geometric results. He presents an interesting argument that the probability of his findings being false, are for all practical purposes, zero. Very briefly, the argument goes as follows. If we try to prove such geometric assertions by using analytic geometry, we wind up with some algebraic relationships which have to be shown to be identities. However, as pointed out by Davis (1977), an algebraic identity can be conclusively established by a single numerical check by using algebraically independent transcendental numbers. Although computers cannot actually operate with transcendental numbers, a series of experiments selecting points at random, achieves much the same result. In other words, if experiment after experiment with randomly selected points reaffirms the same result, the probability of the result being false effectively becomes zero. (Since the property checker of *Cabri* is a trade secret I do not known exactly how it works, but assumes that it probably works along similar lines). Particularly relevant are the following remarks by Grunbaum(1993:8):

> "*Do we start trusting numerical evidence (or other evidence produced by computers) as proofs of mathematics theorems?... if we have no doubt - do we call it a theorem? ... I do think that my assertions about quadrangles and pentagons are theorems ...the mathematical community needs to come to grips with the possibilities of new modes of investigation that have been opened up by computers.* "

Several more examples of the increasing use of the computer, not only as an exploratory tool in different areas, but also as a means of verification, are given in a recent article by Horgan (1993) in the *Scientific American.* Apart from ignoring that there are numerous areas in mathematical research where computers are not (yet?) particularly useful, he then furthermore proceeds from the narrow viewpoint that the **only** function of proof is that of the establishment of mathematical truth and concludes that "*proof is dead*." However, as will be briefly discussed in the next section, a deductive proof may also be useful for many other reasons.

Proof as a means of explanation, discovery and self-realization

Although I have, as described earlier, often achieved confidence[4] in the general validity of a conjecture by its continuous transformation across the screen (or using the property checker), this provides no personally satisfactory *explanation of **why** it may be true.* It merely confirms that it is true, and even though the consideration of more and more examples may increase one's confidence even more, it gives no psychological satisfactory sense of illumination, i.e. an **insight** or **understanding** into how it is the consequence of other familiar results. In fact, it has been my experience that the more convinced I become, the more motivated I also become to find out why it is true. Davis & Hersh (1983:368) similarly state as follows that despite the convincing heuristic evidence in support of the earlier mentioned Riemann Hypothesis, one may still have a burning need for explanation:

> "*It is interesting to ask, in a context such as this, why we still feel the need for a proof .... It seems clear that we want a proof because ... if something is true and we can't deduce it in this way, this is a sign of a lack of understanding on our part. We believe, in other words, that a proof would be a way of understanding **why** the Riemann conjecture is true, which is something more than just knowing from convincing heuristic reasoning that it **is** true.*"

Recently Gale (1990:4) also emphasized that the function of Lanford and other mathematicians' proofs of Feigenbaum's experimental discoveries in fractal geometry was that of explanation and not that of verification at all. Thus, in most cases when the results concerned are intuitively self-evident and/or they are supported by convincing quasi-empirical or computer evidence, the main function/purpose of proof for mathematicians is certainly not that of verification, but rather that of explanation. It is not so much a question of *"making sure"*, but rather a question of trying to *"explain why"*. Of course, the eventual production of an explanatory proof does increase one's confidence, but that is negligible to the understanding it provides.

Admittedly not all proofs are explanatory; sometimes we have to be satisfied with simply a verification proof, but the ideal is generally to try and arrive at some form of satisfactory explanation. I can recall many times where I have spent hours, days and weeks trying to redraft my own non-explanatory proofs into explanatory ones. Of course, one is not always successful, as some results seem to lend themselves more easily to certain classes of proofs that may be considered relatively *non-explanatory*, for instance, indirect, mathematical inductive, analytic geometric and vector proofs (compare Hanna, 1989). It is certainly unthinkable to do without such powerful proof techniques, but for myself I would normally only turn to them as a last resort. Robert Long (1986, 616) writes similarly as follows:

> "*Proofs that yield insight into the relevant concepts are more interesting and valuable to us as researchers and teachers than proofs that merely demonstrate the correctness of a result. We like a proof that brings out what seems to be essential. If the only available proof of a result is one that seems artificial or contrived it acts as an irritant. We keep looking and thinking.*"

Also to Manin (1981:107) and Bell (1976:24), explanation is a criterion for a *"good"* proof when stating respectively that it is *"one which makes us wiser"* and that it is expected *"to convey an insight into* ***why*** *the proposition is true.*"

Another very important function of proof is that of **discovery**. For example, the production of a proof for a result, which identifies its underlying explanatory properties, can sometimes lead to a further unanticipated generalization as was mentioned in Example 3. I have experienced the value of proof as a means of discovery on numerous occasions, thus finding results which would hardly have been likely through more or less "*blind*" experimentation (see for example De Villiers, 1991 & In press).

It is also important to realize that proving something is an intellectual challenge which mathematicians find as appealing as other people may find puzzles or other creative hobbies. It could for instance be compared to the physical challenge of completing an arduous marathon or triathlon. In this sense, proof serves the function of **self-realization** and **fulfillment**. Proof is therefore a testing ground for the intellectual stamina and ingenuity of the mathematician. To paraphrase Mallory's famous comment on his reason for climbing Mount Everest: *"we prove our results because they're there."*

Lastly, it should also be pointed out that a proof may also fulfill several other functions such as *systematization*, *communication*, *memorization* and *algorithmization*. (For more details, consult De Villiers, 1990, Renz, 1981 & Van Asch, 1993).

Concluding comments

Although most students seem to have no further need for conviction having extensively explored geometric conjectures in dynamic geometry environments like *Cabri* or *Sketchpad*, I have nevertheless found it relatively easy to solicit further curiosity by asking them **why** they think a particular result is true; i.e. to challenge them to try and *explain* it. Students then quickly admit that inductive/experimental verification merely confirms; it gives no satisfactory sense of *illumination*; i.e. an insight or understanding into how it is a consequence of other familiar results. Students therefore seem to find it quite satisfactory to then view a deductive argument as an attempt at explanation, rather than verification.

Rather than one-sidedly trying to focuss on proof only as a means of verification in dynamic geometry (which does not make sense to students), the more fundamental function of explanation and discovery ought to be utilized to present proof as a meaningful activity to students (compare Hersh, 1993). This requires that students should be inducted early into the art of problem posing and allowed sufficient opportunity for exploration, conjecturing, refuting, reformulating and explaining as outlined in Figure 12 (compare Chazan, 1990). Dynamic geometry software strongly encourages this kind of thinking as they are not only powerful means of verifying true conjectures, but also extremely valuable in constructing counter-examples for false conjectures.

Movshovitz-Hadar (1988a & b) argues similarly (but not identically) for stimulating presentations of results that solicit the surprise and curiousity of students so that they are susceptable to responsive proofs which leave them with "*an appreciation of the invention, along with a feeling of becoming wiser*".

We should also be quite honest by telling our students that we as mathematicians often prove results simply because of the intellectual challenge involved, and not to try and present a fairy tale of alwyas wanting to obtain "*certainty*". We should also try to give more attention to the communicative aspects of proof by actually negotiating with our students the criteria for acceptable evidence, explanations and/or arguments. Furthermore, as anyone with a bit of experience in actual research will testify, the purely systematization function of proof (i.e. the arrangement of a series of results in a strictly axiomatic-deductive form) comes to the fore only at a very advanced stage, and should therefore be with-held in any introductory course to proof.

In conclusion, the alternative approach to dynamic geometry I am canvassing for is not with the intention of sacrificing any fidelity in mathematics merely for pedagogical expediency, but actually quite the contrary: the encouragement of greater fidelity with respect to the processes of mathematical discovery and the variety of reasons behind proof, some of which were illustrated earlier.

Notes

1The English translation from the French is poor and should actually read "*this property is true in general*". Presumably, the property checker is based on the mathematical theory described in Davis (1977). In this regard, also see the paragraph on *The psychology of mathematical discovery and proof* later on. The property checker of *Cabri* is also able to construct and display a counter-example if the property is not true in general.

2 I used a macro-construction for the construction of the similar triangles on the sides which does not allow for changes to the constructed shape; in other words, their basic *shape* remains *fixed* during the transformation of the base quadrilateral. So strictly speaking, the property checker of *Cabri* in this case only shows that the result is true in general (i.e. for **any** quadrilateral) for these *particular* similar triangles. One could however easily construct other similar triangles on the sides with another macro-construction and repeat the experiment. Although I was sufficiently convinced at this stage, I could have created a dynamic configuration in *Sketchpad* that allows for the dynamic transformation of not only the base quadrilateral, but also the similar triangles on the sides.

3 Similar to the previous case, the respective shapes of the similar rectangles and rhombi on the sides are fixed. (For example, the rectangles were constructed so that their lengths were twice their breadths). Although I was sufficiently convinced at this stage, I could have repeated the experiments with other similar rectangles or rhombi, or used *Sketchpad* to create dynamic configurations which allow for dynamic changes to either the base quadrilateral or the rectangles or rhombi on the sides.

4 I am aware of a few cases with *Cabri* and *Sketchpad* where certain constructions do not work out correctly. For example, in *Cabri* the traditional construction of a common tangent to two circles produces two points on the one circle (instead of only one). Using the checking facility however shows that the linesegments from the center of that circle to the two points are both perpendicular to the tangent, therefore indicating that the two points are actually coincident. (In fact, one cannot distinguish the two points visually in spite of zooming in on them). Similar problems also occasionally arise in *Sketchpad.* In fact, I am aware of a true result in the latter case involving relatively complicated constructions where it is not confirmed (two circles which should end up being tangential to each other are not). However, from a research point of view such a *false negative* is probably not too problematic as one may still suspect the truth of such a result provided the error seems very small, and continues to remain so under transformation. More seriously, would be to find *false positives* (i.e. results continually confirmed by dynamic geometry software, but are actually false). So far, however, I have not yet come across any such results (at least not interesting or significant ones) and have learnt to trust dynamic software to a very large degree.

References

Bell, A.W. (1976). A study of pupils' proof-explanations in mathematical situations. *Educational Studies in Mathematics*, 7, 23-40.

Chazan, D. (1990). Quasi-empirical views of mathematics and mathematics teaching. In Hanna, G. & Winchester, I. (Eds). *Creativity, thought and mathematical proof*. Toronto: OISE.

Davis, P.J. (1977). Proof, completeness, transcendentals and sampling. *Journal Assoc. Comp. Machin.*, 24, 298-310.

Davis, P.J. & Hersh, R. (1983). *The Mathematical Experience.* Great Britain: Pelican Books.

De Villiers, M.D. (1990). The role and function of proof in mathematics. *Pythagoras*, 24, 17-24.

De Villiers, M.D. (1991). Vertical line and point symmetries of differentiable functions. *Int. Journ. Math. Ed. Sci. & Technol.*, 22(4), 621-644.

De Villiers, M.D. (In press). *Some adventures in elementary geometry*.

Gale, D. (1990). Proof as explanation. *The Mathematical Intelligencer***,** 12(1), 4.

Gardner, M. (1981). *Mathematical Circus*. Gt. Britain: Chaucer Press.

Grunbaum, B. (1993). Quadrangles, pentagons, and computers. *Geombinatorics*, 3, 4-9.

Halmos, P. (1984). Mathematics as a creative art. In Campbell, D. & Higgens, J. (1984). *Mathematics: people, problems, results*. Vol. II, 19-29. Belmont: Wadsworth.

Hanna, G. (1989). Proofs that Prove and Proofs that Explain. *Proceedings of the 13th International Conference on the Psychology of Mathematics Education*, Paris, 45-51.

Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics*, 24, 389-399.

Horgan, J. (1993). The death of proof. *Scientific American*, Oct, 74-82.

Kline, M. (1982). *Mathematics. The Loss of Certainty.* Oxford: Oxford University Press.

Kolata, G.B. (1989). A program that makes conjectures - mathematicians meet computerized ideas. *New York Times*, 18 June 1989.

Long, R.L. (1986). Remarks on the history and philosophy of mathematics. *The American Mathematical Monthly*, 93, 609-619.

Manin, Y.I. (1981). A Digression on Proof. *The Two-year College Mathematics Journal*, 12(2), 104-107.

Movshovitz-Hadar, N. (1988a). Stimulating presentations of theorems followed by responsive proofs. *For the Learning of Mathematics*, 8(2), 12-19;30.

Movshovitz-Hadar, N. (1988b). School mathematics theorems - An endless source of surprise. *For the Learning of Mathematics*, 8(3), 34-40.

Pollak, H.O. (1984). The effects of technology on the mathematics curriculum. In Olivier, A.I. (Ed). *The Australian Experience: Impressions of ICME 5*. Centrahil: AMESA.

Polya, G. (1954). *Mathematics and Plausible Reasoning. Induction and Analogy in Mathematics.* Vol 1. Princeton: Princeton University Press.

Renz, P. (1981). Mathematical proof: What it is and what it ought to be. *The Two-Year College Mathematics Journal*. 12(2), 83-103.

Van Asch, A.G. (1993). To prove, why and how? *Int. J. Math. Educ. Sci. Technol***.**, 24(2), 301-313.

Yaglom, I.M. (1962). *Geometric Transformations I.* Washington, DC: Mathematical Association of America.