Paper given at 9th National Congress of AMESA, 30 June-4 July 2003, Cape Town, In Proceedings, pp. 174-185.

THE VALUE OF EXPERIMENTATION IN MATHEMATICS

Michael de Villiers, University of Durban-Westville
profmd@mweb.co.za
http://mzone.mweb.co.za/residents/profmd/homepage.html

INTRODUCTION

The purpose of this paper is to investigate what the role of experimentation is in mathematics, reflecting on some historical examples and some from my own mathematical experience. With experimentation is meant here all non-deductive methods including intuitive, inductive or analogical reasoning. In other words, it is specifically employed when:

(i) mathematical conjectures and/or statements are numerically or visually evaluated, by means of special cases, accurate geometric construction and measurement, etc.

(ii) conjectures, generalisations or conclusions are made on the basis of intuition, analogy or experience obtained through any of the preceding experimental methods, etc.

The most important functions of experimentation (in no specific order of importance) can be distinguished as follows, though these are quite often closely linked, as will be illustrated in the following discussion and examples:

* *conjecturing* (looking for an inductive pattern, generalisation, etc.)
* *verification* (obtaining certainty about the truth or validity of a statement or conjecture)
* *global refutation* (disproving a false statement by generating a counter-example)
* *heuristic refutation* (reformulating, refining or polishing a true statement by means of local counter-examples)
* *understanding* (the meaning of a proposition, concept or definition or assisting with the discovery of a proof).

CONJECTURING

The history of mathematics is literally replete with hundreds of examples where conjectures were made merely on the basis of intuition, numerical investigation and/or construction and measurement. A good example is the famous Prime Number theorem which was first formulated in about 1792 by Gauss. Using numerical evidence obtained from counting prime numbers and using logarithms, Gauss discovered that the number of prime numbers smaller or equal to a number *n*, is always approximately $\frac{n}{\log n}$, and the approximation improved as *n* increased. While several mathematicians already actively used the conjecture at the beginning of the nineteenth century to explore different properties of prime numbers, a partial proof of it was only given in 1850 by Chebychev. Although the conjecture was generally accepted as proved from 1859 onwards when Riemann published a more complete proof, there were still some gaps in his proof that were only filled in later, independently of each other, by Hadamard and De La Vallee' Poussin in 1896.

This example also shows that mathematicians may sometimes, even in the absence of rigorous proofs, accept certain inductively confirmed conjectures as "*theorems*", especially if they are in an important field of research.

Polya (1954:3) similarly strongly emphasises the importance of experimentation in the discovery or invention of new mathematics, and quotes one of the most productive mathematicians of all time, namely, Leonhard Euler, in this regard as follows:

> "*As we must refer the numbers to the pure intellect alone, we can hardly understand how observations and quasi-experiments can be of use in investigating the nature of numbers. Yet in fact ... the properties of the numbers known today have been mostly discovered by observation, and discovered long before their truth has been verified by rigid demonstrations.*"

From my own (though relatively mundane) mathematical experience, I can also give several different examples where my own initial conjectures flowed almost entirely from prior experimental experience. For example, a couple of years ago I made the following two dual conjectures (see De Villiers, 1991), while investigating the symmetry of several different graphs and their associated derivatives using graphing software:

(1) A differentiable function in the plane is reflective symmetric around a vertical axis $y = a$ if and only if its derivative is point symmetric around the point $(a; 0)$;

(2) A differentiable function in the plane is point symmetric around a point (*a; b*) if and only if its derivative is reflective symmetric around the axis $y = a$.

The arrival of the modern computer, as an extremely powerful tool for experimental exploration, has in the past few decades also revolutionised mathematical research in several areas, resulting in many new and exciting results.

One of the main advantages of computer exploration of topics is that it provides powerful visual images and intuitions that can contribute to a person's growing mathematical understanding of that particular research area. Furthermore, the computer provides a unique opportunity for the researcher to formulate a great number of conjectures and to immediately test them by only varying a few parameters of a particular situation.

Even traditional Euclidean geometry is experiencing an exciting revival, in no small part due to the recent development of dynamic geometry software such as *Cabri*, *Sketchpad* and *Cinderella*. In fact, Philip Davies (1995) predicts as a consequence a particularly rosy future for resurgence in triangle geometry research. Recently Adrian Oldknow (1995) for example used *Sketchpad* to discover the hitherto unknown result that the Soddy center, incenter and Gergonne point of a triangle are collinear (amongst other interesting related results). Similarly, I recently experimentally discovered a generalisation of Neuberg's theorem (De Villiers, 2002), and rediscovered a beautiful generalisation of the nine-point circle of a triangle to a nine-point ellipse (conic), as well as an associated generalisation of the Euler line.

Much is often made of the crucial role of "*intuition*" in mathematical discovery and invention. Perhaps most significantly from an educational point of view is that most authors strongly emphasise that intuition is dependent on "*experience*" rather than just an innate, natural ability. In other words, it mostly develops from the regular handling, exploration and manipulation of mathematical objects. Such experience obviously does not only refer to formal logical manipulation, but also to the experimental exploration of objects.

VERIFICATION/CONVICTION

Contrary to the traditional belief amongst many mathematics teachers that only proof provides certainty for the mathematician, mathematicians are often convinced of the truth of their results (usually on the basis of experimental evidence), long before they have proofs. Indeed, as argued in De Villiers (1990), conviction is often a prerequisite for looking for a proof. If one were uncertain about a result one would rather look for a counter-example, not a proof. However, one needs to be reasonably convinced about the truth of a result before sitting down and possibly spending a considerable amount of time generating a proof.

In real mathematical research, personal conviction usually depends on a combination of experimentation and the existence of a logical (but not necessarily entirely rigorous) proof. In fact, a very high level of conviction may sometimes be reached even in the absence of a proof. For instance, Davis & Hersh (1983) present extremely convincing "*heuristic evidence*" in support of the still unproved Riemann Hypothesis in terms of numerical evidence and a statistical model developed by Good & Churchhouse in 1968, and conclude that this evidence is "*so strong that it carries conviction even without rigorous proof.*"

That this kind of experimental conviction often precedes and motivates a proof is borne out in the history of mathematics, i.e. by the frequent heuristic precedence of results over arguments, of theorems over proofs. For example, Gauss is reputed to have complained: *"I have had my results for a long time, but I do not yet know how I am to (deductively) arrive at them"*. Bernhard Riemann also exclaimed in some frustration: "*If only I had the theorems! Then I should find the proofs easily enough*.".

Furthermore, experimental evidence not only frequently plays a role in the initial formulation of a conjecture, but quite often also in continuing efforts to prove a particular result. Let us consider a very simple example. Since an isosceles trapezoid has (at least) one opposite pair of parallel sides, as well as equal diagonals, it seems reasonable to conjecture that these might be sufficient conditions for defining an isosceles trapezoid. However, suppose one does not fairly quickly come up with a proof (and the reader is invited to try and prove this before reading further), one would naturally start wondering whether it is indeed true. Perhaps the conjecture is false and one is trying to prove something that is not true!

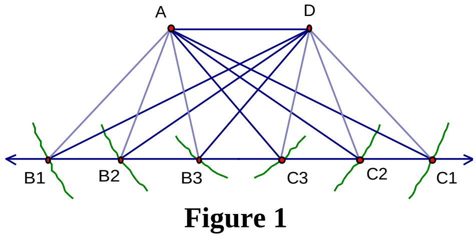

Figure 1

However, by accurately (or even roughly) drawing a line segment *AD* and a line parallel to it, and then equal diagonals *AC* and *DB*, as shown in Figure 1 to test it, one can intuitively see, even without measurement, that opposite sides $AB_n = DC_n$, irrespective of how or where the diagonals $AC_n = DB_n$ are drawn. Even better, one could do the construction in a dynamic geometry environment to gain an even higher level of confidence. Now armed with the knowledge that a counter-example cannot be constructed and that it is definitely true, one can now with renewed confidence resume looking for a proof.

GLOBAL REFUTATION

In everyday life people often use a kind of fuzzy logic, i.e. believing certain things to be true if it is true most of the time, simply ignoring the occasional cases when they aren't true. Unlike everyday life, however, mathematical theorems can have no exceptions, and only one counter-example is sufficient to disprove a mathematical proposition. With "*global refutation*" is meant here the production of a logical counter-example that meets the conditions of a statement, but refutes the conclusion, and thus the validity of the statement.

Generally in mathematics, global counter-examples are also produced by experimental testing, and usually not by deductive reasoning. Consider the following false conjecture from elementary geometry: "*a quadrilateral with perpendicular diagonals is a kite*". To construct a counter-example for this statement it is only necessary to check experimentally whether sufficient information is provided for the construction of a kite. If one now constructs two perpendicular diagonals and let the various segments have arbitrary lengths as shown in Figure 2, one easily finds that the constructed figure is not necessarily a kite. Similarly, one would not use deduction to construct counter-examples for conjectures such *as "a quadrilateral with equal diagonals is an isosceles trapezoid" or "6x - 1 is a prime number for all x = 1, 2, 3, etc.*", but experimental testing.

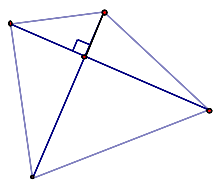

Figure 2

There are many examples from the history of mathematics that clearly illustrate how experimental testing generated counter-examples, though sometimes taking many years. For example, in the fifth century BC Chinese mathematicians already made the conjecture that if $2^n - 2$ is divisible by *n* then *n* is a prime number. If this was true it would have been valuable for determining the primality of a number, as then one would only have to carry out the division of $2^n - 2$ by *n*. Approaching the conjecture inductively, one finds: $2^3 - 2$, $2^5 - 2$, $2^7 - 2$ are divisible by the primes 3, 5 and 7, but $2^4 - 2$, $2^6 - 2$, $2^8 - 2$ are not divisible by the composite numbers 4, 6 and 8.

It turns out that experimental investigation supports the conjecture up to $2^{340} - 2$ (a very large number indeed!). In all these cases $2^n - 2$ is divisible by *n* when *n* is prime, and not divisible by *n* when it is composite. However, this conjecture was disproved only in 1819 when it was found that $2^{341} - 2$ is divisible by 341, but 341 is not prime since $341 = 11 \times 31$.

HEURISTIC REFUTATION

With reference to the historical development of the Euler-Descartes theorem, Lakatos (1983) argues that proof is not a mechanical and infallible procedure for obtaining truth and certainty in mathematics. Instead he views proof as a collection of explanations, justifications and interpretations which become increasingly more acceptable with the continued absence of counter-examples.

An important distinction often not picked up by a naive, casual or mathematically inexperienced reader of Lakatos is that between global counter-examples, and "*local*" or "*heuristic*" counter-examples. Whereas the former, like those in the previous paragraph completely disprove a statement, the latter challenge perhaps only one step in a logical argument or merely aspects of the domain of validity of the proposition. Heuristic counter-examples are mostly not strictly logical counter-examples, since they are after all not inconsistent with the conjecture in its intended interpretation, but are heuristic, since they spur the growth of knowledge.

Additional experimental testing of an already proved statement may therefore sometimes assist one in finding subtle gaps in a proof or in the formulation of a theorem. Let me illustrate this with a very simple personal example in relation to the earlier mentioned dual conjectures regarding vertical line and point symmetric functions, namely:

(1) A differentiable function in the plane is reflective symmetric around a vertical axis $y = a$ if and only if its derivative is point symmetric around the point (*a; 0*);

(2) A differentiable function in the plane is point symmetric around a point (*a; b*) if and only if its derivative is reflective symmetric around the axis $y = a$.

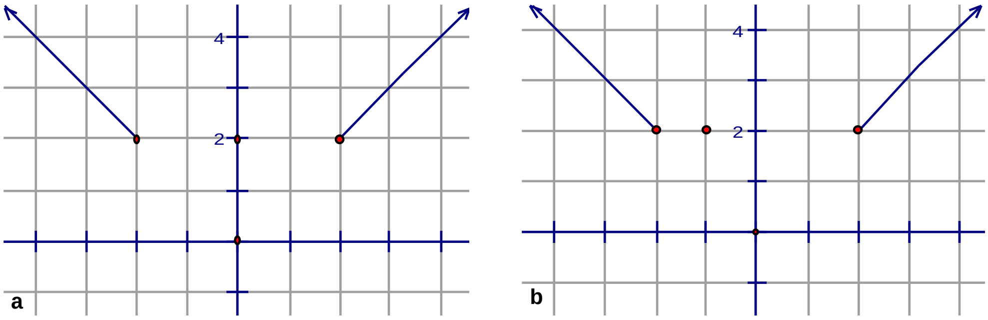

Figure 3

After providing proofs for these two conjectures, both geometric and analytic, I was somewhat later idly considering the discontinuous function defined by $y = -x$ for $x < -2; y = 2$ for $x = 0$ and $y = x$ for $x > 2$ as shown in Figure 3a. Clearly the derivative ($\frac{dy}{dx} = -1$ for $x < -2$ and $\frac{dy}{dx} = 1$ for $x > 2$) is point symmetric around (0;0). However, consider now the function $y = -x$ for $x < -2; y = 2$ for $x = -1$ and $y = x$ for $x > 2$ shown in Figure 3b. Note that the reflective symmetry is now destroyed, but its derivative remains point symmetric around (0;0).

So here I had a counter-example to the converse part of Theorem 1! But how could this be? Had I not already proved these theorems?

However, proceeding to check my proofs, I quickly realised that the proofs only referred to the differentiable parts of the function, and that these would be symmetrical, even if the function as a whole is not symmetrical. All that was required, therefore, was to reformulate the two theorems more precisely as follows:

(1) The differentiable parts of a function in the plane are reflective symmetric around a vertical axis $y = a$ if and only if the derivatives of these parts are point symmetric around the point (*a; 0*);

(2) The differentiable parts of a function in the plane are point symmetric around a point (*a; b*) if and only if the derivatives of these parts are reflective symmetric around the axis $y = a$.

UNDERSTANDING

As already mentioned in the previous section, the experimental investigation and evaluation of already proved results, can sometimes lead to new perspectives and a deeper understanding or extension of earlier concepts and definitions. Indeed, it is a common practice among mathematicians while reading someone else's mathematical paper to look at special or limiting cases to help them unpack and understand not only the results better, but also the proofs. During our own research, it can also sometimes assist us to more rigorously define our intuitive concepts, which can in turn lead to new investigations in hitherto uncharted directions.

For example, based on some research described in De Villiers (1989), I was made to realise that the concept of "*interior*" angles of complex polygons, with sides criss-crossing each other, was not so intuitively obvious as it may seem. Indeed, the "*interior*" angle of a crossed polygon could actually sometimes lie "*outside*"!

This counter-intuitive observation would probably not have been possible without experimental investigation. It also helped me to rethink carefully the meaning of interior angles in such cases, eventually coming up with a consistent, workable definition.

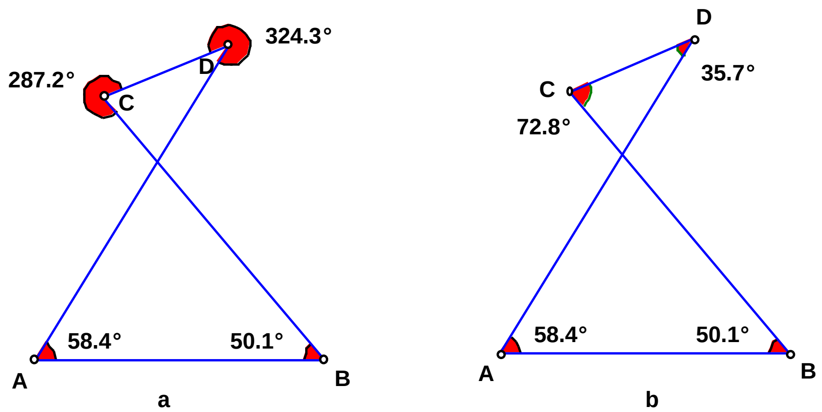

Figure 4

Using this definition, I next made another surprising, counter-intuitive discovery, namely, that the interior angle sum of a crossed quadrilateral is always 720° (see Figure 4a). Indeed, this specific example can be used as a simple, but authentic illustration of the method of heuristic refutation, and works extremely well at the high school level, as well as with mathematics teachers (see De Villiers, 2003).

In fact, almost without exception, the first reaction of most people when confronted with the type of figure shown in Figure 4b, and asked to determine its interior angle sum, is that of "*monster-barring*"; i.e. a blunt rejection of such a figure as a quadrilateral. A very common response is to argue that it can't be a quadrilateral, since its angle sum is not 360°. Another response is sometimes to say we should just add the two opposite angles where the two sides *BC* and *AD* intersect to ensure that the angle sum is still 360° (despite 6 angles now being involved!).

Clearly what is at stake here is what we choose to understand under the concepts "*quadrilateral*", "*vertex*" and "*interior angle*", and not the validity of the result that the angle sum of convex and concave quadrilaterals is 360° - that is undisputed. In fact, mathematically the situation can easily be resolved either by defining quadrilaterals in such a way as to exclude crossed quadrilaterals or simply to explicitly state in the formulation of the theorem that it only applies to simple closed quadrilaterals (convex and concave).

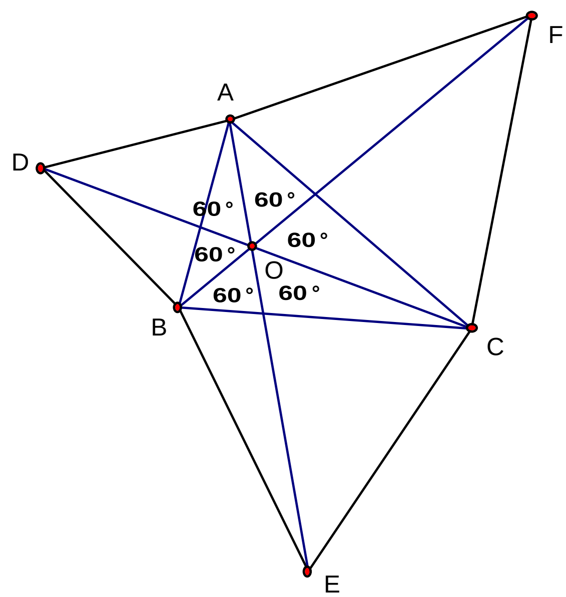

Figure 5

However, experimental investigation can also sometimes contribute to the discovery of a hidden clue or underlying structure of a problem leading eventually to the construction/invention of a proof. For example, consider Figure 5 showing equilateral triangles on the sides of an arbitrary triangle, and that the lines *DC*, *EA* and *FB* are concurrent (in the so-called Fermat-Torricelli point). Noting by dragging with dynamic geometry, or otherwise, that the six angles surrounding point *O* are all equal, can assist one to recognise *FAOB*, *DBOC* and *ECOA* as cyclic quadrilaterals, setting one well on the way to constructing a synthetic proof.

EXPERIMENTAL-DEDUCTIVE INTERPLAY

Undoubtedly, in everyday research mathematics, quasi-empiricism and deduction complement each other, rather than oppose. Generally, our mathematical certainty does not rest exclusively on either logico-deductive methods or quasi-empiricism, but usually on a healthy combination of both. Intuitive thought and experimental experience broaden and enrich, and do not only stimulate deductive reflection, but can contribute to the critical quality of such deductive reflection by the provision of heuristic counter-examples. Intuitive, informal (experimental) mathematics is therefore an integral part of genuine mathematics.

The limitations of intuition and experimental investigation should also not be forgotten. Even George Polya (1954:v), famous advocate for heuristic, informal mathematics, warns that intuitive, experimental thinking on its own can be "*hazardous"* and "*controversial*". A good example is that of Cauchy who had the popular intuition of his time that the continuity of a function implied its differentiability. However, at the end of the 19th century, Weierstrass stunned the mathematical community by producing a continuous function that was not differentiable in any point!

Presumably inspired by Fermat's Last Theorem, Euler also conjectured that there were no integer solutions to the following equation:

$$x^4 + y^4 + z^4 = w^4$$

For two hundred years nobody could find a proof nor could anyone disprove it by providing a counter-example. First calculation by hand and then years of computer sifting failed to provide a counter-example, namely, a set of integer solutions. Indeed many mathematicians started believing it was true, and that it was probably only going to be a matter of time before someone came up with a proof. However, then in 1988 Naom Elkies from Harvard University discovered the following counter-example:

$$2682440^4 + 15365639^4 + 18796760^4 = 20615673^4.$$

An even more spectacular example of the danger of reliance on only experimental evidence is the following from Rotman (1998:3):

> Investigate whether S(n) = $991n^2 + 1$ is a perfect square or not. What do you notice? Can you prove your observations?

Random numerical investigation for several n is likely to show that $991n^2 + 1$ is not a perfect square. In fact, the statement is true for **all** n until:

n = 12 055 735 790 331 359 447 442 538 767

$\approx 1.2 \times 10^{29}$

Despite having so much evidence **-** far more evidence than there have been days on earth **-** the conjecture turns out to be FALSE!

Indeed, nobody today can really be considered mathematically educated or literate, if that person is not aware of the insufficiency of only experimental evidence to guarantee truth in mathematics, no matter however convincing that evidence may seem.

Besides not providing sufficient certainty, experimental evidence as pointed out earlier, also seldom provides satisfactory explanations. As a result, deductive proofs are invaluable to explain, justify, and systematise our mathematical results. In addition, proving results may in turn lead to further generalisations or spawn research in different directions.

In a similar vein, the research mathematician Gian-Carlo Rota (1997:190) pointed out, regarding the recent proof of Fermat's Last Theorem, that the value of the proof goes far beyond that of mere verification of the result:

> "*The actual value of what Wiles and his collaborators did is far greater than the mere proof of a whimsical conjecture. The point of the proof of Fermat's last theorem is to open up new possibilities for mathematics. ... The value of Wiles's proof lies not in what it proves, but in what it opens up, in what it makes possible.*"

CONCLUSION

It is simply intellectually dishonest to pretend in the classroom that conviction only comes from deductive reasoning or that adult mathematicians never experimentally investigate conjectures and already proved results. Why deny students the opportunity to explore conjectures and results experimentally when we as adult mathematicians quite often indulge in such activities in our own research? Even though such testing by students may not produce any heuristic counter-examples, it may still help students better understand the propositional meaning of a theorem.

ACKNOWLEDGEMENT

This research was partially funded by a National Research Foundation (NRF) grant from the Spatial Orientation & Spatial Insight (SOSI) Project, GUN No. 2050502, coordinated by Prof. Dirk Wessels (UNISA), Prof. Hercules Nieuwoudt (PUCHE) and Prof. Michael de Villiers (UDW).

REFERENCES

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

Davies, P.J. (1995). The rise, fall, and possible transfiguration of triangle geometry. *American Mathematical Monthly*, 102(3), 204-214.

Davis, P. J. & Hersh, R. (1983). *The Mathematical Experience*. Harmondsworth: Penguin Books.

De Villiers, M. (1989). From "TO POLY" to generalized poly-figures and their classification: a learning experience. *Int. J. Math. Ed. Sci. Technol*, 20(4), 585-603, August 1989.

De Villiers, M. (1990). The role and function of proof in mathematics. *Pythagoras*, 24, 17-24. . (A copy of this paper can be directly downloaded from http://mzone.mweb.co.za/residents/profmd/proofa.pdf).

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

De Villiers, M. (2002). From nested Miquel Triangles to Miquel distances. *Mathematical Gazette*, 86(507), Nov., 390-395.

De Villiers, M. (2003). *Rethinking Proof with Geometer's Sketchpad 4*. Emeryville: Key Curriculum Press, USA.

Lakatos, I. (1983). *Proofs and Refutations.* 4th Reprint. Cambridge: Cambridge University Press.

Oldknow, A. (1995). Computer aided research into triangle geometry. *The Mathematical Gazette*, 79(485), 263-274.

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

Rota, G-C. (1997). The Phenomenology of Mathematical Beauty. *Synthese*, 111, 171-182.

Rotman, J. (1998). *Journey into Mathematics: An Introduction to Proofs*. NY: Prentice Hall.