The Role and Function of Quasi-empirical Methods in Mathematics[1]

Michael de Villiers
University of KwaZulu-Natal

Abstract: This article examines the role and function of so-called *quasi-empirical methods* in mathematics, with reference to some historical examples and some examples from my own personal mathematical experience, in order to provide a conceptual frame of reference for educational practice. The following functions are identified, illustrated, and discussed: *conjecturing*, *verification*, *global refutation*, *heuristic refutation*, and *understanding*. After some fundamental limitations of quasi-empirical methods have been pointed out, it is argued that, in genuine mathematical practice, quasi-empirical methods and more logically rigorous methods complement each other. The challenge for curriculum designers is, therefore, to develop meaningful activities that not only illustrate the above functions of quasi-empirical methods but also accurately reflect an authentic view of the complex, interrelated nature of quasi-empiricism and deductive reasoning.

Sommaire exécutif: Cet article analyse le rôle et la fonction des méthodes dites « quasi-empiriques » en mathématiques, par le biais de certains exemples historiques et d'autres provenant de ma propre expérience, dans le but de fournir un cadre de référence conceptuel l'enseignement. Les fonctions identifiées, illustrées et analysées sont les suivantes :

- *la conjecture* (recherche par induction, généralisation, analogie, etc.)
- *la vérification* (tentative d'obtenir des certitudes sur la vérité ou la validité d'une affirmation ou d'une hypothèse) ;
- *la réfutation globale* (démonstration du fait qu'une affirmation est fausse grâce à la génération d'un contre-exemple) ;
- *la réfutation heuristique* (reformulation, affinement ou perfectionnement d'une affirmation essentiellement vraie par le biais de contre-exemples ponctuels) ;
- *la compréhension* (compréhension d'un théorème, d'un concept, d'une définition ou d'une démonstration, ou encore contribution à la découverte d'une preuve ou à la formulation précise d'une définition).

Nous nous intéressons en particulier à l'utilisation de plus en plus courante de l'informatique pour explorer les différents sujets, car l'ordinateur fournit des images visuelles et d'autres stimuli qui alimentent les intuitions susceptibles de contribuer à une meilleure compréhension d'un secteur de recherche donné en mathématiques. Nous soulignons la distinction importante qui existe entre la réfutation globale et la réfutation heuristique. En effet, si la première vise à démontrer la fausseté des résultats mathématiques, la seconde contribue à raffiner et à reformuler aussi bien les résultats que leur démonstration. Après une brève analyse des avantages que présentent les méthodes quasi-empiriques dans chacune des catégories citées plus haut, nous donnons des exemples qui soulignent les limites de ces méthodes pour ce qui est des certitudes (certains résultats, par exemple, résistent à de nombreuses épreuves avant de céder). De plus, il est rare que les méthodes quasi-empiriques servent à approfondir le niveau de compréhension (par exemple à comprendre *pourquoi* les résultats sont vrais), et il est également rare qu'elles contribuent à une systématisation des mathématiques chez les étudiants (par exemple qu'elles servent à établir des liens, etc.)

Dans la pratique des mathématiques, nous estimons donc que, loin de s'opposer les unes aux autres, les méthodes quasi-empiriques et les méthodes plus rigoureuses sur le plan de la logique se complètent. Le défi à relever dans la mise au point des curriculums consiste à créer des activités significatives capables non seulement d'illustrer les fonctions citées plus haut, mais également de fournir une vision authentique de la complexité des liens qui caractérisent les raisonnements quasi-empiriques et les raisonnements déductifs.

Introduction

Something often lacking in our mathematical education at school or university level is providing students with a sense of how new results can or could be discovered or invented. Quite often, after theorems and their proofs have been carefully presented, students are just given exercises with riders of the type 'Prove that …'. This caricature of mathematics can easily create the false impression that mathematics is only a systematic, deductive science. However, as George Polya (on many occasions), Imre Lakatos (1983), and many others have pointed out, mathematics in the making is often an experimental, inductive science.

But what are these experimental methods, and more importantly, what role do they play in the discovery and creation of new mathematics? How can these be utilized in the mathematics classroom, not only to provide a more authentic view of mathematics, but also to make mathematics a more meaningful, exciting experience for students?

The main purpose of this article is to address the first question; namely, to investigate the role of experimentation in mathematics, although aspects related to the second question will occasionally be interwoven into the discussion. It is hoped that this will provide a useful conceptual frame of reference for designing, as well as for evaluating, learning activities and curricula. This analysis will involve reflecting on some historical examples and some examples from my own mathematical experience. It is important for educational practitioners to explore what mathematicians actually do in their practice and not base their practice on some distorted illusion about what they think mathematicians do. Moreover, students who excel at exploration and conjecturing have a substantial contribution to make to mathematics and the mathematical sciences but are currently being side-tracked away from these studies by the narrow misrepresentation, in their classrooms, of what mathematics really is.

Instead of the term *experimental* methods, I will mostly use the more general term *quasi-empirical* methods. This term is borrowed from Lakatos (1983), since the objects in mathematics, though largely abstract and imaginary, can be subjected to empirical testing much as scientific theories are. Quasi-empiricism will, therefore, refer here to all non-deductive methods involving experimental, intuitive, inductive, or analogical reasoning. In other words, it is specifically employed when

- mathematical conjectures and/or statements are evaluated in a non-deductive manner; for example, numerically, visually, graphically, through special cases, through construction and measurement, diagrammatically, through physical embodiments, kinaesthetically, by analogy, and so on;
- conjectures, generalizations, or conclusions are made on the basis of intuition or experience obtained through any of the preceding quasi-empirical methods.

Though laying claim to neither completeness nor originality, I have found the following model for the role of quasi-empirical methods very useful in my own curriculum design and educational practice. This model of some important functions of quasi-empirical methods is presented now (in no specific order of importance) and is discussed in more detail further on:

- *conjecturing*—looking for an inductive pattern, generalization, analogy, and so on
- *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 an essentially true statement by means of local counter-examples

- *understanding*—comprehending a theorem, concept, definition, or proof or assisting the discovery of a proof or precise formulation of a definition

This model will now be illustrated by examples more typical of *pure* mathematics than of *applied* mathematics (or science). This means not that the model does not apply to the mathematical modelling of real-world problems, but that the different functions of experimentation may acquire some subtly different interpretations. However, a more complete analysis of the role of quasi-empirical methods in modelling contexts is best left for another article.

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 geometric construction and measurement. A good example is the famous Prime Number Theorem, which was first formulated, in about 1792, by Gauss.

Table 1: Possible process of reasoning for arriving at Gauss's prime number

Number (n)	Primes less than n $\pi(n)$	$n/\pi(n)$
10	4	2.5
100	25	4.0
1,000	168	6.0
10,000	1,229	8.1
100,000	9,592	10.4
1,000,000	78,498	12.7
10,000,000	664,579	15.0
100,000,000	5,761,455	17.4
1,000,000,000	50,846,534	19.7

Although it is not known precisely how Gauss arrived at this result, Davis and Hersh (1983, pp. 213–214) suggest the following plausible scenario. If one tabulates some numbers of primes _*n*) less than or equal to the corresponding numbers *n*, as shown in Table 1, then it seems natural to compute the ratio *n*/ _*n*). It then seems likely that Gauss noticed that, as one moves from one power of 10 to the next, this ratio increases by roughly 2.3. Since $ln(10) = 2.30258 \ldots$, it is reasonable to conjecture that the number of prime numbers smaller or equal to a number *n*, is always approximately $n/\ln n$ (ln is the natural logarithm with base *e*).

While several mathematicians at the start of the nineteenth century actively used this conjecture 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 Vallée Poussin in 1896 (see, e.g., Davis & Hersh, 1983; Kramer, 1981; Hanna, 1983).

As Hanna points out (1983, p. 73), this historical 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. Another case in point is the still unproved Riemann hypothesis, which was also discovered essentially by induction but has been used extensively in the twentieth century to discover new results in number theory by examining some of its logical consequences (on the assumption that it is true). Neubrand (1989, p. 2), for example, describes the Riemann hypothesis as a fundamental 'driving force in the dynamic development of mathematics, as e.g. Euclid's Parallel Postulate did some centuries before: one looks for consequences which are maybe provable by elementary means, or which may eventually yield a contradiction to the hypothesis.'

Goldbach's still unproved conjecture, namely, that any even number greater than 2 can be written as the sum of two primes, was already formulated by 1742 and was also, apparently, based exclusively on experimental evidence. For example, $4 = 2 + 2$; $6 = 3 + 3$; $8 = 3 + 5$, $10 = 3 + 7$; $12 = 5 + 7$; and so on. Another unresolved conjecture is that there are infinitely many twin primes (primes that differ by 2) such as (3, 5) or (11, 13). Though no proof for either of these two conjectures has yet been found, most mathematicians are convinced that they are likely to be true.

Eves (1969, pp. 167–168) similarly points out that many of the results of Euclidean geometry can actually be traced back to the building, constructing, surveying, and astronomical techniques of the ancient Egyptians and Babylonians, which essentially rested only on empirical grounds. It is certainly not inconceivable that several geometric conjectures were probably first made purely on the basis of visual observation, trail and error, and/or other experimental evidence.

Lakatos (1983), citing T.L. Heath, has pointed out that the ancient 'Greeks did not think much of propositions, which they happened to hit upon in the deductive direction without having previously guessed them. They called them porisms, corollaries, incidental results springing from the proof of a theorem or the solution of a problem, results not directly sought, but appearing, as it were, by chance, without any additional labour, and constituting, as Proclus says, a sort of windfall (*ermaion*) or bonus (*kerdos*)' (p. 9).[1]

Polya (1954) also very strongly emphasizes the importance of quasi-empirical exploration in the discovery or invention of new mathematics and quotes one of the most productive mathematicians of all time—namely, Leonhard Euler—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' (p. 3).

The arrival of the modern computer, as an extremely powerful tool for experimental exploration, has, in the past few decades, also revolutionized mathematical research in several areas, resulting in many new and exciting results. Indeed, one could say that the computer has become an indispensable tool for making conjectures and exploring complicated, abstract relationships in modern mathematics. Hofstadter (1985) prosaically describes some of these developments in relation to the mathematical modelling of turbulence and other chaotic phenomena, as follows:

> Probably the main reason these ideas are only now being discovered is that the style of exploration is entirely modern; it is a kind of experimental mathematics, in which the digital computer plays the role of Magellan's ship, the astronomer's telescope, and the physicist's accelerator. Just as ships, telescopes, and accelerators must be ever larger, more powerful, and more expensive in order to probe even more hidden regions of nature, so one would need computers of even greater size, speed and accuracy in order to explore the remoter regions of mathematical space (p. 365).

One of the main advantages of computer exploration of topics is that it provides powerful visual images and stimuli for intuitions that can contribute to a growing mathematical understanding of a given research area. Furthermore, the computer provides a unique opportunity for the researcher to formulate a great number of conjectures and to test them immediately, by varying only a few parameters of a particular situation. In fact, Hofstadter (pp. 366–369) argues, it simply would not have been possible, with traditional, non-computer-based research, to have arrived so quickly and easily at such a rich, coherent body of new results in many areas of mathematics.

Not surprisingly, in 1991, a successful new quarterly journal, *Experimental Mathematics*, was established, whose main mission is not only finished theorems and proofs but also an exploration of the experimental ways in which results have been reached—in other words, to also display the dynamic interaction between theory and experimentation in research mathematics (see Epstein & Levy, 1995).

Even traditional Euclidean geometry research is experiencing an exciting revival, due in no small part 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, 1996), for example, used *Sketchpad* to discover the hitherto unknown result that the Soddy centre, incentre, and Gergonne point of a triangle are collinear (amongst other interesting related results).

A few years ago, the geometer June Lester (1997) also used *Sketchpad* to experimentally discover that the two Fermat points, the nine-point centre, and the circumcentre of a triangle are always concyclic (on a circle which is now appropriately called the Lester circle). Similarly, I recently experimentally discovered generalizations of the Fermat-Torricelli point of a triangle (see De Villiers, 1999) and of Neuberg's theorem (see De Villiers, 2002), respectively.

At the heart of making such new conjectures lies the ability to look for and be unafraid to ask 'what if …?' questions from many different perspectives; in other words, to become a good *problem poser.* A major difference between problem solving and problem-posing is that the former tends to be convergent, whereas the latter is divergent. Good problem posing, therefore, requires a willingness to try out new ideas and to explore new avenues of thought: to guess, experiment, and test.

This is an important habit of mind we should actively strive to encourage in our students at all levels. By continually nurturing such a questioning, inquiring habit, not only is one stimulating divergent problem posing skills in one's students, but some of these questions may occasionally lead to entirely new and exciting results. From my own experience, I have also found that my students are far more highly motivated to solve problems they've discovered and formulated for themselves. It seems that the personal ownership and involvement stemming from one's own discoveries are very effective in stimulating genuine intellectual curiosity.

Reasoning by means of analogy to arrive at new conjectures is, unfortunately, also something that is not demonstrated frequently enough to students at high school and university level. For example, from Viviani's theorem (that the sum of the three distances from a point inside an equilateral triangle to its sides is constant), one can easily conjecture that the same might be true for any regular polygon. Or going into three dimensions, one can easily conjecture that the sum of the four distances from a point inside a regular tetrahedron to its faces is also constant.[2] Polya (1954, 1968, 1981) gives many examples that can suitably be adapted for mathematics teaching at various levels.

Much is often made of the crucial role of *intuition* in mathematical discovery and invention. However, most authors strongly emphasize that intuition is dependent more on *experience* than on some innate, natural ability. In other words, it mostly develops from the regular, prolonged handling, exploration, and manipulation of mathematical objects and ideas (cf. Davis & Hersh, 1983, pp. 391–392; Epstein & Levy, 1995; Volmink, 1990, p. 7; Wilder, 1967). Such experience, obviously, includes not only formal logical manipulation but also the quasi-empirical exploration of those objects and ideas, often over days, months, or years. This has significant educational implications for designing curricula and learning materials.

Intuition is often used simply as a clumsy shorthand for a deep and critical process. In particular, if processing is not in language, or in the numbers or symbols of an algorithm or construction, it is considered *intuition*. For example, visual processing, with several discrete steps and transformations, may occur long before an idea is first put into words or symbols. Einstein's work with thought experiments, including his caution to not put things into words or symbols too soon, is, in fact, typical of this process. This is certainly a developed ability, which ought to be encouraged, modelled, and assessed in mathematics education.

At some point, our students also need to learn that quasi-empirical investigation is not always a prerequisite for making conjectures and arriving at solutions to problems. The following is a good example that might be used: Consider finding the total number of tennis matches played in an elimination tournament if there are n players.

Most students will approach the problem inductively, by looking at cases $n = 2, 3, 4$, and so on, and then generalize. However, a more astute student may, by just thinking carefully about the situation, quickly realize that the total number of matches must be $n - 1$, since there can only be one winner and there must, therefore, be $n - 1$ matches to eliminate the other $n - 1$ players.

Suppose no student sees the crucial, initial conceptualization (looking at the losses rather than the rounds), and they all proceed to solve the problem the hard way? In such a case, their attention can still be directed to wondering why the answer is one less than the number of players and whether this is a signal that they have missed the essence of the situation and an opportunity to solve it more elegantly.

Through such activities, students could learn that reflective, logical thought may, indeed, be more powerful and more appropriate than immediately embarking on a quasi-empirical search for pattern.[3] De Jager (1990, p. 12) gives some good examples of these and argues that, generally, one should start searching inductively or experimentally only as a last resort.

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 quasi-empirical 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 look for a counter-example, not a proof. One needs to be reasonably convinced about the truth of a result before sitting down and spending what may turn out to be a considerable amount of time generating a proof.

In real mathematical research, personal conviction usually depends on a combination of quasi-empirical 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, Polya (1954) quotes Leonhard Euler on an important discovery he made in the algebra of the real numbers:

> It suffices to undertake this multiplication and to continue it as far as it is deemed proper to become convinced of the truth of this series. Yet I have no other evidence for this, except a long induction which I have carried out so far that I cannot in any way doubt the law governing the formation of these terms and their exponents. I have long searched in vain for a rigorous demonstration of the equation between the series and the above infinite product … and I have proposed the same question to some of my friends with whose ability in these matters I am familiar, but all have agreed with me on the truth of this transformation of the product into a series, without being able to unearth any clue of a demonstration (p. 96).

Similarly, the geometer Branko Grünbaum (1993) used the computer program *Mathematica* to explore and verify some geometric results; his concluding comments are highly relevant: ‘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 my assertions are theorems … the mathematical community needs to come to grips with new modes of investigation that have been opened up by computers’ (p. 8).

That this kind of quasi-empirical conviction often precedes and motivates a proof is borne out by the history of mathematics; that is, 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' (cited in Arber, 1954, p. 47). Bernhard Riemann also exclaimed, in some frustration, 'If only I had the theorems! Then I should find the proofs easily enough' (cited in Hölder, 1924, p. 487). Paul Halmos (1984) underscores this idea when he describes his own practice as a mathematician as follows: 'The mathematician at work … arranges and rearranges his ideas, and he 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' (p. 23).

The practice of first evaluating an unknown conjecture by the consideration of specific cases is probably as old as mathematics itself and is still actively utilized in modern-day research. Neubrand (1989), for example, writes as follows about the proof of Bieberbach's conjecture (1916) (now De Branges' theorem [1984]): 'As in many other cases, in this example mathematicians first started with the consideration of special cases, restricted cases, etc., in order to convince themselves of the possibility of the validity of the conjecture' (p. 4). Furthermore, quasi-empirical evidence frequently plays a role not only 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!

However, by accurately (or even roughly) drawing a line segment AD and a line parallel to it, and then drawing equal diagonals AC and DB, as shown in Figure 1 to test, 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 counterexample cannot be constructed and therefore that any trapezoid with equal diagonals is definitely an isosceles trapezoid, one can now with renewed confidence resume looking for a proof.

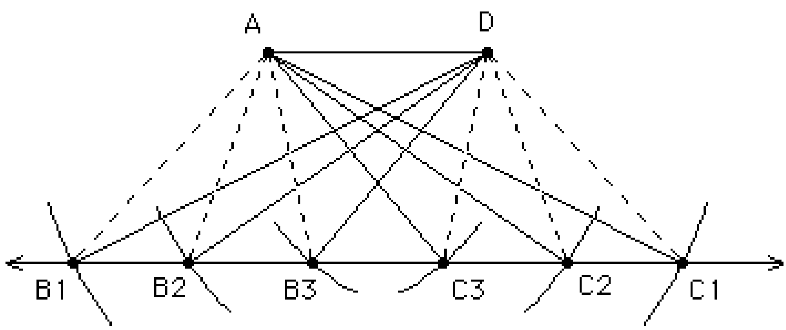

Figure 1

It should be emphasized that quasi-empirical conviction obtained via computer investigation has not, as Horgan claims (1993), made proof obsolete. Indeed, modern-day computer investigations frequently stimulate further curiosity and thus, also, a need for proof as a means of explanation—that is, in order to clarify and better understand why things are true. A particular case in point is the recent proofs by Lanford and other mathematicians of Feigenbaum's experimental discoveries in fractal geometry (see, e.g., Gale, 1990).

Global refutation

In everyday life, people often use a kind of fuzzy logic; that is, believing certain things to be true if they are true most of the time, simply ignoring (often conveniently) the occasional cases when they are not true. Unlike in everyday life, however, mathematical theorems can have no exceptions, and only one counter-example is sufficient to disprove a mathematical proposition. By *global refutation* is meant here, therefore, the production of a logical counter-example that meets the conditions of a statement but refutes the conclusion, and thus the general validity, of the statement.

Global counter-examples at the elementary level are usually produced by quasi-empirical testing and not by deductive reasoning. Consider the following conjecture from elementary geometry, made by students in a class considering the defining properties of a kite: 'A quadrilateral with perpendicular diagonals is a kite.' To construct a counter-example for this false statement, it is only necessary to check quasi-empirically whether sufficient information is provided for the construction of a kite. If one now constructs two perpendicular diagonals and lets the various segments have arbitrary lengths, as shown in Figure 2, one easily finds that the constructed figure is not necessarily a kite.[4]

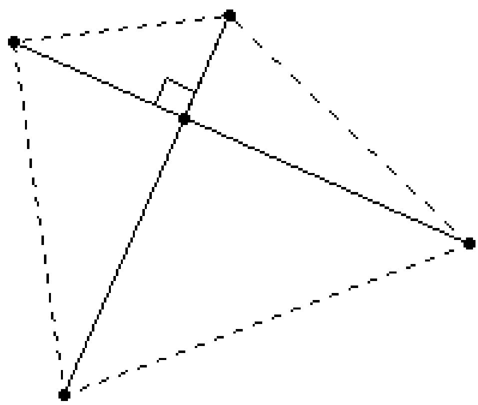

Figure 2

Similarly, to construct counter-examples for conjectures such as 'A quadrilateral with equal diagonals is an isosceles trapezoid'[5] or '$6x$ - 1 is a prime number for all x = 1, 2, 3 …', one would use not deduction but quasi-empirical testing.

When one is dealing with a totally unknown mathematical conjecture, therefore, it is often prudent to test it quasi-empirically first, as this testing serves (at least) one of the following purposes:

1 the construction of a counter-example if it is false
2 the attainment of a reasonable amount of certainty (conviction), which then encourages one to start looking for a proof

There are many examples from the history of mathematics that clearly illustrate how quasi-empirical testing generates counter-examples. Sometimes, it takes many years before a conjecture is refuted. For example, in the fifth century BC, Chinese mathematicians had already conjectured that, if $2^n - 2$ is divisible by n, then n is a prime number (see Kramer, 1981, p. 514). If this were true, it would be valuable for determining the primality of a number, as one would then only have to carry out the division of $2^n - 2$ by n. Approaching the conjecture inductively, one finds that $2^3 - 2$, $2^5 - 2$, $2^7 - 2$ 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 quasi-empirical 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, the conjecture was disproved in 1819, when it was found that $2^{341} - 2$ is divisible by 341, but 341 is not prime, since $341 = 11 \times 31$.

A more contemporary example is Lord Kelvin's long-standing conjecture, from about 1850, that the optimal partition of space into equal volumes with the minimum total surface is obtained by warping the tiling of space by truncated octahedra. Everyone seemed satisfied with Kelvin's solution; and it was believed that it was only a matter of time before a proof of optimality was produced. However, using a computer program, *Surface Evolver*, the physicists Weaire and Phelan produced, in 1994, a space partition of equal volumes with a considerably smaller surface area than that in Kelvin's solution (Epstein & Levy, 1995; Hales, 2000). It is not known whether this is the best possible solution, so the Kelvin problem is still open.

However, not all counter-examples are constructed by quasi-empirical testing. For example, since 41 is clearly a factor of $n^2 - n + 41$, when $n = 41$, one might easily notice, without any quasi-empirical substitution, that this provides an immediate counter-example to the conjecture that $n^2 - n + 41$ is always prime for $n = 1, 2, 3 \ldots$ In a historical reconstruction, Waterhouse (1994) also suggests that Gauss must have used substantial theoretical argumentation to arrive at the counter-example he gave in 1807 to a conjecture of Sophie Germain's. An even more spectacular example is Mertens' conjecture. Despite the fact that, by 1963, there was already computer-supported evidence for the conjecture's holding true for all n up to 10 million, Odlyzko and Te Riele gave an existential proof for the existence of a counter-example to it in 1984 (without constructing an actual counter-example).

Quasi-empirical testing is also useful for identifying incorrect assumptions in otherwise completely valid reasoning. Consider, for example, the following simple, but instructive, example from De Villiers (1996) that I have often used successfully with students—a seemingly convincing proof of a reasonable conjecture.

Conjecture

A quadrilateral with one pair of opposite sides equal and one pair of opposite angles equal is a parallelogram.

Proof

Consider Figure 3, with $AB = CD$ and $\angle BAD = \angle DCB$. Draw BX+AD and DY+BC and join B and D. Then, it is easy first to prove triangles AXB and CYD congruent, and then triangles BXD and DYB congruent. Therefore, $AX = CY$ and $XD = YB$ from which we have $AD = AX + XD = CY + YB = CB$. In other words, $ABCD$ is a parallelogram (opposite sides are equal).

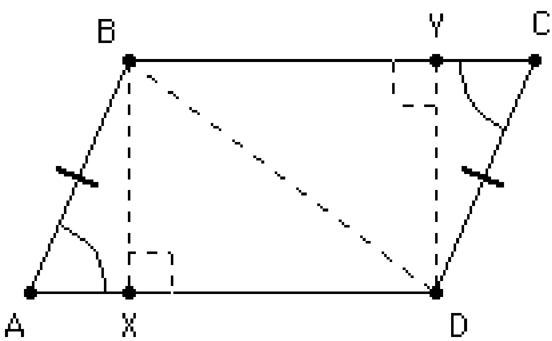

Figure 3

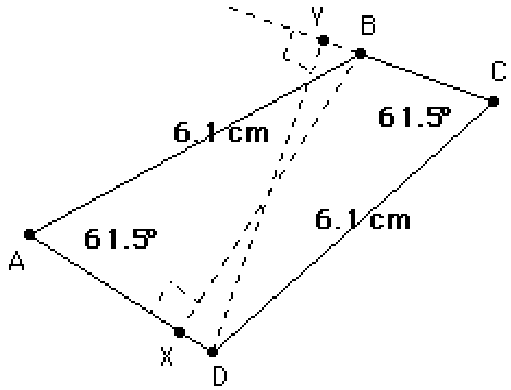

Figure 4

My experience with students has invariably been that they find no fault with this argument and uncritically accept it as true, not bothering to look for a counter-example at all. However, the argument is based on the false assumption that the constructed perpendiculars always fall inside the quadrilateral. It, therefore, comes as quite an instructive surprise when they are asked to construct a figure that meets those requirements, using dynamic software; some then find counter-examples, such as that shown in Figure 4, where *Y* falls on *CB* extended.[6]

I have found this simple example useful in showing students how important it sometimes is to check deductive arguments for faulty assumptions empirically, by accurate construction and measurement. In general, dynamic geometry software is particularly useful for investigations in geometry, as a configuration can easily and quickly be dragged into many different configurations to check the general validity of an assumption. Although it is sometimes jokingly claimed that geometry is the art of drawing accurate conclusions from inaccurate figures, an example like that above can serve to caution students not to draw too inaccurately!

In fact, many ingenious geometric paradoxes, such as *all triangles are isosceles*, can arise by virtue of construction errors or mistaken assumptions in diagrams (cf. Bradis, Minkovskii, & Kharcheva, 1959; Northrop, 1980; Movshovitz-Hadar & Webb, 1998). Not only is unravelling paradoxes by pinpointing precise reasoning errors or mistaken assumptions educationally instructive; but, as Kleiner & Movshovitz-Hadar (1994) point out, historically, paradoxes have also contributed to the evolution of many parts of mathematics. Paradoxes should not, therefore, be considered just an amusing sideline but should be incorporated and integrated into mathematics education wherever appropriate.

Quasi-empirical evaluation is also valuable in identifying an erroneous proof of a true result. For example, some years ago, I was investigating possible generalizations and variations of Napoleon's theorem and was fortunate to discover the following interesting result experimentally, with the aid of dynamic geometry (see De Villiers & Meyer, 1995),

If similar triangles *ABD*, *EBC*, and *AFC* are erected on the sides of any triangle *ABC*, as shown in Figure 5, their respective incentres *G*, *H*, and *I* form a triangle with $\angle IGH = ½(\angle DAB + \angle DBA)$, $\angle GHI = ½(\angle EBC + \angle ECB)$, and $\angle HIG = ½(\angle FCA + \angle FAC)$.

I clearly recall how I sat in a boring meeting, attempting to while away my time more productively by looking for a proof for this particular result. After drawing a variety of auxiliary lines on my rough sketches, and several unsuccessful proof attempts, I suddenly saw a proof, and everything seemed to fall almost magically into place like a jigsaw puzzle. I literally had to contain myself from exclaiming with excitement. Fortunately, the meeting was just about at an end, and I could escape quickly and rush off to my office to check my proof with dynamic geometry on my computer. Alas, to my great dismay, I found that the relative positions of certain crucial points and lines to the auxiliary construction that I had drawn roughly, were actually incorrect, and so, unfortunately, my 'proof' was invalidated! It was back to the drawing board, and only much later was a satisfactory proof found.

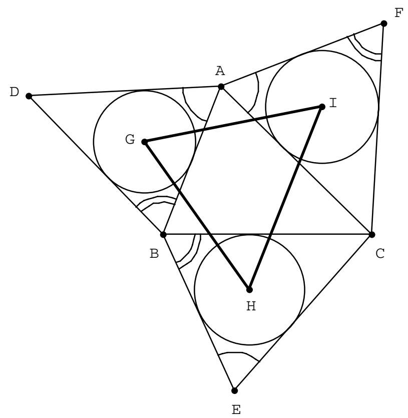

Figure 5

Heuristic refutation

Although mathematics is not an empirical science, it grows and develops, according to Lakatos (1983), in a way very similar to that of the natural sciences; that is, as a consequence of the quasi-empirical testing of theorems, concepts, definitions, and so on. New counter-examples necessitate the re-examination of old proofs, and new proofs are created accordingly.

Lakatos analysed the history of Euler's theorem[7] for polyhedra and dramatized it within a fictional classroom context. After first stating in 1750, without proof, that for polyhedra such as the tetrahedron, octahedron, and so on, $V - E + F = 2$, where V, E, and F are the numbers of vertices, edges, and faces, respectively, Euler eventually gave a proof in 1752. His proof was followed by more rigorous proofs, in the nineteenth century, by Legendre, Cauchy, Gergonne, Rothe, and Steiner. Nevertheless, there continued to be some exceptions or *monsters*, such as Kepler's star dodecahedron (see Figure 6), for which Euler's formula was not valid. Only towards the latter part of the nineteenth century did topologists finally manage to develop a completely satisfactory proof, based on a more precise, general definition of polyhedra and on generalizing the formula to $V - E + F = 2 - 2g$ (where g is the *genus* or number of *holes* in the polyhedron—see Grünbaum & Shephard, 1994; Hilton & Pedersen, 1994).

Lakatos attributes the inordinately long delay in resolving the Euler theorem to the leading mathematicians of the time's not having realized that immediately after the appearance of heuristic counter-examples, they ought to have closely examined the proofs to identify the guilty lemmas. Instead, they tended to treat the heuristic counter-examples in a stereotypical way; that is, simply ignoring them or rejecting them as monsters and excluding them by definition. According to Lakatos (1983, p. 137–139), this *monster-barring* process was a direct consequence of the dominant view that deductive proof was always infallible and, therefore, that formal proofs were above scrutiny and unquestionable.

From a Lakatosian viewpoint, therefore, it is useful, by means of quasi-empirical exploration, to test not only unproved conjectures but also results already proven deductively. Such testing ought also to be encouraged, rather than suppressed, among our students, as it may give new perspectives for further research or contribute to the refinement and/or reformulation of earlier proofs, definitions, and concepts. The Lakatosian view, therefore, contrasts strongly with a traditional, rationalist view like that of Fischbein (1982) that a 'formal proof offers an absolute guarantee to a mathematical statement. Even a single practical check is superfluous' (p. 17).

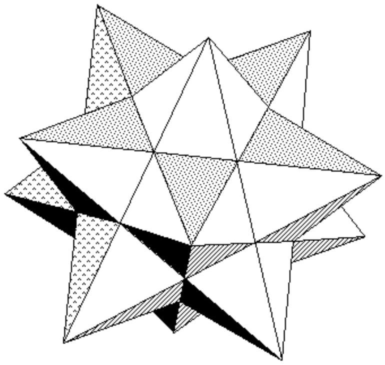

Figure 6

However, 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 section on global refutation, 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. Most heuristic counter-examples, therefore, are not strictly logical counter-examples, since they are not inconsistent with the conjecture in its intended interpretation, but are heuristic in that they spur the growth and refinement of knowledge.

Note, therefore, that in general, a heuristic counter-example only requires some reformulation of the theorem or its proof and usually leaves the original theorem relatively intact. In other words, after all is said and done, the original conjecture (theorem) is usually still valid and true, not disproved at all, though perhaps modified, refined, and much better understood. An excellent example, possibly accessible at senior high school level if transformed into a learning activity, but probably more appropriate at undergraduate level, is described fully in De Villiers (2000).

Generally, mathematical theorems (and theories) exhibit a permanence often denied to proofs, which may change according to the prevailing rigour of the time. For example, even though Euclid did not prove (or even state as an axiom) the Jordan Curve Theorem, namely, that a closed curve like a circle or triangle has an inside and an outside, this *hole* in Euclidean geometry does not destroy or invalidate his work at all.

Recent tendencies to derive or develop a fallibilist philosophy of mathematics education—tendencies usually justified from a Lakatosian perspective—are, however, unfortunate. For example, Gila Hanna (1995, 1997) has pointed out that there are many historical cases where mathematical development has been radically different from the heuristic refutation described by Lakatos. Surely, as mathematics educators and mathematicians, we ought to know the danger of over-generalizing from only a few historical cases!

Nevertheless, radical fallibilist views appear to have become a dominant, fashionable ideology in current mathematics education (see, e.g., Ernest, 1991; Borba & Skovsmose, 1997). The claim that all mathematics is potentially flawed and open to correction is certainly problematic. It would appear that an underlying, implicit assumption is that this Lakatosian process of proof and heuristic refutation can, in principle, carry on indefinitely. However, this assumption is really not supported by history, and in fact, the majority of our rich mathematical inheritance, at least at school and undergraduate level, can be regarded as 'rock bottom,' as Davis and Hersh point out (1983): 'Once a proof is "accepted," the results of the proof are regarded as true (with very high probability). It may take generations to detect an error in a proof. If a theorem is widely known and used, its proof frequently studied, if alternative proofs are invented, if it has applications and generalizations and is analogous to known results in related areas, then it comes to be regarded as 'rock bottom'! (p. 354).

Undeniably, our mathematical results, once proven correctly (though perhaps subject to later revision), are still the most certain of all human knowledge. I am definitely (several orders of magnitude!) more certain of the universal validity of the Pythagorean theorem in the plane, than I am about whether the sun will rise tomorrow. Even if the earth were suddenly to be destroyed tomorrow, this would not alter the theorem's validity elsewhere in the universe. It is misleading, therefore, and not in the best educational interests of our students, to deny the existence of this extremely high level of certainty in mathematics.

Understanding

As already mentioned in the previous section, quasi-empirical 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 common practice among mathematicians, while reading someone else's mathematical article, to look at special or limiting cases to help them unpack and better understand not only the results but also the proofs.

As mathematicians, we need examples to ensure we know what the words, symbols, and notations of a proof mean. Ideally, we use multiple examples, from different contexts, if the abstract theorem applies in multiple places. This is true for both practitioners and learners. Apart from deepening understanding, the examples also add certainty to a complex proof. To practising mathematicians, reading a proof and doing some examples is not irrelevant. Fischbein's claim (1982) that a 'formal proof offers an absolute guarantee to a mathematical statement. Even a single practical check is superfluous' (p. 17), therefore, does not really hold true in practice. In fact, a practical check is something one ought to actively encourage one's students to do. The match between the check and the proof makes both the proof and the working of the example clearer and more convincing. One source of possible 'error' in current mathematics is the occasional subtle switch in the meaning of terms, symbols, and so on, between one mathematical article and another, a switch that usually only becomes apparent upon examining a few special cases.

Quasi-empirical testing can sometimes assist us in defining our intuitive concepts more rigorously, and this sometimes, in turn, leads to new investigations in hitherto uncharted directions. For example, once I was attempting to generalize the interior angle sum formula $(n - 2) \times 180°$ for simple closed polygons to more complex polygons, with sides criss-crossing each other. In the process, I rediscovered, for myself, that the concept of the *interior* angles of some complex polygons was not quite so intuitively obvious (see De Villiers, 1989). Indeed, I found that the interior angle of some crossed polygons 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 and eventually come up with a consistent, workable definition. 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 7a). Indeed, this specific example can be used as a simple, but authentic, illustration of a heuristic counter-example and works extremely well at the high school level, as well as with mathematics teachers (see De Villiers, 2003, pp. 40–44).

In this learning activity, students who already know and have established the theorem that the sum of the interior angles of any quadrilateral is 360° are confronted with the type of figure shown in Figure 7b. Almost without exception, the first reaction of most students is typically that of *monster-barring* in defence of the theorem; that is, a blunt rejection of such a figure as a quadrilateral. Most commonly, it is argued that it can't be a quadrilateral, since its angle sum is not 360°. To this observation, some students sometimes respond by saying that we could add the two opposite angles where the two sides *BC* and *AD* intersect to ensure that the angle sum remains 360° (conveniently ignoring that they are now involving 6 angles!).

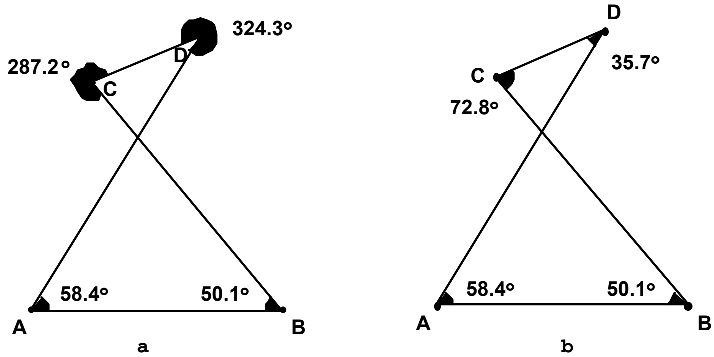

Figure 7

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. Indeed, mathematically, the situation can easily be resolved either by defining quadrilaterals in such a way as to exclude crossed quadrilaterals or simply by explicitly stating, in the formulation of the theorem, that it only applies to simple, closed quadrilaterals (convex and concave).

It is, in fact, typical of refutation by heuristic counter-example to stimulate argument about the precise meaning of the concepts involved and to propose and criticize different definitions of these, until consensus is achieved. For example, after the discovery of a counter-example to the Euler theorem for polyhedra, the characters in *Proofs and Refutations* vehemently argue about whether to accept or reject the counter-example (see Lakatos, 1983, p. 16).

Quasi-empirical investigation can also sometimes contribute to the discovery of a hidden clue or underlying structure of a problem that can eventually lead to the construction of a proof. For example, consider Figure 8, 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). Now noting, perhaps by dragging with dynamic geometry, that the six angles surrounding point *O* are all equal can assist one to recognize *FAOC*, *DBOA*, and *ECOB* as cyclic quadrilaterals (since in each case the exterior angle at *O* is equal to the opposite interior angle), setting one, possibly, on the way to constructing a synthetic proof.

Polya (1968) similarly argues that analogy and experimentation can contribute greatly to discovering and understanding proofs: '[A]nalogy and particular cases can be helpful both in finding

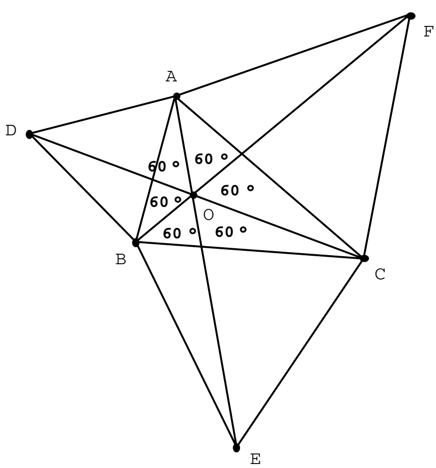

Figure 8

and in understanding mathematical demonstrations. The general plan, or considerable parts, of a proof may be suggested or clarified by analogy. Particular cases may suggest a proof; on the other hand, we may test an already formulated proof by how it works in a familiar or critical particular case' (p. 168). An instructive example of how investigation with dynamic geometry assisted me in the eventual construction of a proof for an interesting generalization is given in De Villiers (1997, pp. 18–20). After several unsuccessful proof attempts, I was completely stumped and baffled. I simply could not seem to find a way of tackling the problem. I wracked my brains, looked up a few related results to see if I could find a starting point, tried drawing some auxiliary lines, and so on. All in vain. Then, while dynamically manipulating the configuration for the umpteenth time, I suddenly noticed two angles, visually appearing always to remain equal. Holding my breath, I nervously measured these two angles and found that they were, indeed, exactly equal and remained so under dragging! That was the breakthrough, the concrete clue, I had needed! And soon it enabled me to construct a proof that also explained why the result was true.

Quasi-empirical–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 on a healthy combination of both. Intuitive thought and quasi-empirical experience broaden and enrich: Not only do they stimulate deductive reflection; they can contribute to the critical quality of such deductive reflection by providing heuristic counter-examples. Intuitive, informal (quasi-empirical) mathematics is, therefore, an integral part of genuine mathematics (cf. Wittmann, 1981, p. 396).

Schoenfeld (1986) describes as follows how students used both quasi-empiricism and deduction to solve a problem:

> [T]he most interesting aspect of this problem session is that it demonstrates the dynamic interplay between empiricism and deduction during the problem-solving process. Contributions both from empirical explorations and from deductive proofs were essential to the solution … Had the class not embarked on empirical investigations … the class would have run out of ideas and failed in its attempt to solve the problem. On the other hand, an empirical approach by itself was insufficient. (pp. 245–249)

The limitations of intuition and quasi-empirical investigation should, however, not be forgotten. Even George Polya (1954, p. v), famous advocate for heuristic, quasi-empirical mathematics, warns that such thinking, on its own, can be 'hazardous' and 'controversial.' An excellent 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 nineteenth 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 (see Singh, 1998, p. 178): $x^4 + y^4 + z^4 = w^4$. 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 that Euler's conjecture was true and that it was only a matter of time before someone came up with a proof. Then, in 1988, Naom Elkies from Harvard University discovered the following counter-example: $2{,}682{,}440^4 + 15{,}365{,}639^4 + 18{,}796{,}760^4 = 20{,}615{,}673^4$.

An even more spectacular example of the danger of reliance only on quasi-empirical evidence is the following investigation, adapted from Rotman (1998, p. 3), which I regularly use with my own mathematics education students:

991n^2 + 1

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

991n^2 + 1 Systematic or random calculator or computer investigation for several n strongly suggests that ~~991_n) + 1~~ is never a perfect square. Even though some are already practising teachers, my students are usually easily convinced about the truth of this conjecture, particularly after some random testing of the conjecture with a wide range of numbers, including some very large ones on a TI-92 (see Figure 9).

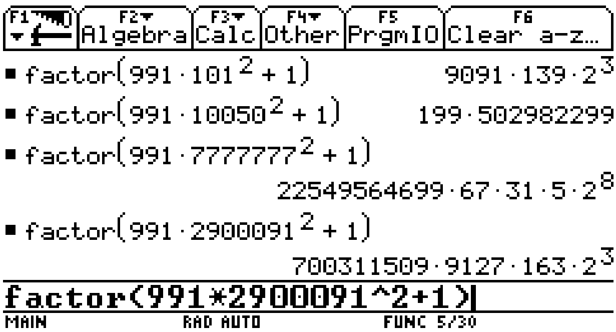

Figure 9

I then challenge my students to produce proofs, and on occasion, some have come up with some ingeniously devised 'proofs.' Even after I have pointed out the errors in their arguments, they usually remain fully confident of the truth of the conjecture; so it then comes as a great shock (and, therefore, an excellent learning experience!) when it is pointed out to them that the statement is true for all n only until:[8] n = 12,055,735,790,331,359,447,442,538,767 _ 1.2×10^{29}

Despite there being so much evidence to support it—in fact far more evidence than there have been days on earth (about 7.3×10^{12} days)—the conjecture turns out to be false!

This example also highlights a fundamental difference between mathematics and science that our students, at all levels, ought to be made more aware of; namely, that science in general is ultimately based on empirical assumptions (even though deduction plays an important part in the mathematical sciences). We simply assume that the *regularities* we observe, like a stone falling to the ground or the sun rising every day, will always hold. What evidence do we have for this assumption? None, except that, as far as we know, it has been correct for millions of years. Still, that the sun will rise every day is simply an empirical assumption and there is no mathematical proof that it will always be correct.

Nobody, today, can really be considered mathematically educated or literate, if he or she is not aware of the insufficiency of quasi-empirical evidence to guarantee truth in mathematics, no matter how convincing that evidence may seem. This should, therefore, be a crucial aim of any mathematics curriculum at the high school level and higher, and students ought to be led to experience proof as an empowering, liberating, and highly intellectually satisfying endeavour (cf. Hanna, 1977). It is, therefore, unfortunate that, in certain parts of the world, there has been a marked decline in the teaching of proof at school level and, in some cases, a virtually complete removal of proof from the curriculum, as happened in the United Kingdom from about the mid-1980s to the mid-1990s. This decline can, perhaps, in part be attributed to the increased dominance of a radical fallibilist viewpoint in mathematics education, presumably influenced by a superficial interpretation of the statement by Lakatos (1983, p. 143) that proof is 'the worst enemy of independent and critical thought.' However, what Lakatos was criticizing here was not proof per se, but the traditional direct teaching of pre-existing proofs, which, without a proper balance of conjecturing and adequate quasi-empirical exploration, is, indeed, the enemy of independent and critical thought in the classroom.

Besides not providing sufficient certainty, quasi-empirical evidence also seldom provides satisfactory explanations; that is, insight into why something is true in mathematics. In other words, a drawback of quasi-empirical investigation is that it does not tell us how a result relates to other results or how it fits into the general mathematical landscape. It is largely for this reason that Rav (1999) has emphasized that proofs, in many respects, rather than theorems, are the really valuable bearers of mathematical knowledge. As a result, students need to experience the value of deductive proofs in explaining, understanding, and systematizing mathematical results. In addition, specific learning activities need to be devised to show them sometimes how proving results may lead to further generalizations or spawn investigations in different directions, as described by Rav (1999).

The research mathematician Gian-Carlo Rota (1997) has similarly pointed out, regarding the recent proof of Fermat's Last Theorem, that the value of the proof goes far beyond that of the 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' (p. 190). Students ought, also, to be more regularly exposed to multiple quasi-empirical approaches to and proofs of the same result. Often, mathematicians have delighted in giving additional proofs of their own or other people's theorems. Clearly, the value here is largely in gaining multiple perspectives; developing a deeper, richer understanding; or opening up a whole range of possible analogies, connections, specializations, and generalizations that can be further explored. Moreover, if the only role of proof were certainty, there would definitely be no interest in alternate proofs (or further quasi-empirical investigations) of existing results. There would also be no great preference for elegant proofs—'proofs from the book'—to quote Erdös (see Aigner & Ziegler, 2004). However, a research question for mathematics educators is: How can the interest of students in alternate proofs be stimulated and maintained if they have already found one of their own?

In order for both proof and quasi-empirical methods to make sense, it seems that students will have to develop a sense of mathematics as a coherent system, where there is some kind of relationship among assertions, and complex properties can be accounted for in terms of more basic ones. Students need to develop an appreciation of structure and infinity and to become sceptical, through the use of numerous examples that show that things may not be as they appear.

The spontaneous, naive approach of students is usually just to check assertions by looking at all possible cases or at least enough of them to resolve their doubt. However, students need to appreciate that this is not possible in mathematical situations where there may be infinitely many cases or there are factors that do not kick in for simple cases.

For example, students might be asked to look at the familiar example of the number of regions into which a disc can be subdivided by all possible lines joining pairs of n points on its circumference (no three lines concurrent). Students are likely to conjecture by looking at cases $n = 2$, 3 and 4 that the maximum number of regions is given by 2^{n-1}. Though this formula still holds for $n = 5$, it breaks down when $n = 6$, when the maximum number of regions is only 31.[9]

Despite the highly convincing power of modern calculator and computer technology, ways of also nurturing a critical scepticism within this context need to be explored. Students need to learn to guard against over-reliance on technology and be aware that, if they are not careful, false conclusions may still be drawn. Despite their fantastic computational power, computers can only calculate discretely and can never verify any conjecture for an infinite number of cases. For example, Srinivasan (2002) and De Villiers (2003, pp. 73, 85) give examples of possible learning activities, within the contexts of calculus and dynamic geometry, respectively, that could be used with great effect to develop some caution in students.

A significant difficulty in formulating proofs, at whatever level, is to know what should be included and what can be left out. It is generally too onerous to make a proof completely comprehensive. The decision is based on one's own sense of what is self-evident, what is significant, and what can be expected of the audience. History suggests that investigators initially tend to take certain elements for granted until paradoxes or counter-examples emerge or some astute individual has the wit to wonder what can possibly go wrong. The resulting increase in the precision of definitions and tightening of arguments comes from a combination of experience, imagination, improved use of language, and refinement of mathematical sense. The trick is to help the young experience this without confusing them or bogging them down with issues that cannot be intelligible to them.

Concluding comments

As strongly argued here, quasi-empirical methods are an integral, indispensable part of mathematical thinking. However, authentic, exciting, and meaningful ways of incorporating them in mathematics education need to be explored if we are, indeed, serious about providing students with a deeper, more holistic insight into the nature of our subject. More specifically, the enormous challenge facing teachers and curriculum designers is to illustrate and develop some understanding and appreciation of the identified functions of quasi-empirical methods; for example, conjecturing, verifying, global and heuristic refutation, and understanding.[10]

For example, students need to develop some insight into how new mathematics is (or can be) discovered and created, and therefore, they require substantial exposure to the complementary roles played by quasi-empirical and deductive thinking in this process. Students also need to realize that mathematicians are often not just convinced (to a high degree) by quasi-empirical evidence but also motivated by it to search for a deductive proof (which, in such cases, usually fulfils more an explanatory than a verification role.)

Instead of mostly providing students with rather artificial 'to-prove' results, a more natural, stimulating environment needs to be created, in which they can come up with some of their own conjectures. Not only would this give students greater personal ownership over mathematics, but their student-generated conjectures might or might not be true and so might provide excellent opportunities for disproof, by means of producing global counter-examples, through quasi-empirical testing.

However, the greatest challenge facing mathematics teachers and curriculum designers is to devise mathematically authentic activities to illustrate the process of heuristic refutation (as well as to help students achieve deeper understanding and clarity by quasi-empirical means). It is time to move beyond Lakatos, Gödel, and the great, unsolved philosophical problems in mathematics. We should transcend the vacuous rhetoric to create more practical guidelines governing the selection of tasks, the preparation needed for students to embark on them and express their ideas cogently, and the approach necessary so that the work of the students leads to robust mathematical competence and deeper understanding. Unfortunately, most curriculum materials currently in circulation seem at best to contain rather mathematically dubious examples, in this regard.

One possible strategy already mentioned is the active promotion of greater student participation (cf. Luthuli, 1996). For example, student-generated conjectures, proofs, and definitions are likely to evolve through a few cycles, being reformulated a few times as a result of the provision of appropriate heuristic counter-examples. This might be enhanced by encouraging students to critique each other's conjectures and proofs and to look critically, through quasi-empirical testing, for counter-examples, as well as for hidden assumptions in proofs (cf. Chazan, 1990).

An additional strategy might be to infuse the mathematics curriculum with carefully chosen historical examples that illustrate the evolution of some important theorems, as well as of false conjectures. Some possible examples have been given in this article, but numerous other historical examples could be used to exemplify the complementary roles played by quasi-empirical methods and deductive reasoning.

Significant support for the inclusion of quasi-empirical methods in the mathematics curriculum at all levels also comes from learning theories such as the Van Hiele theory, which argues that visualization and informal reasoning are important prerequisites for developing higher, abstract reasoning (see Burger & Shaughnessy, 1986; Fuys, Geddes, & Tischler, 1988).

It is quite heartening that computer programs such as dynamic geometry, statistical packages, spreadsheets, function graphing, computer algebra, and so on are already being used by many teachers around the world to encourage and stimulate quasi-empirical thinking via student explorations. It is necessary, though, to push beyond the view of computing technology as a mere tactical hook to attract or simply engage students and start emphasizing its role in the core practice of mathematics, at all levels, from novice and young to expert and mature. Not only does computing software provide powerful means of confirming true conjectures; it is also extremely valuable in constructing counter-examples for false conjectures.

Basically, the role of computers is to make quasi-empirical exploration both easier to execute and more fecund, so that the type of playing around that was once the preserve of only the most persistent or imaginative is now widely accessible; computers, judiciously used, make possible the democratization of mathematical processes. At the level of the classroom, the important question is how to encourage students to work purposefully rather than aimlessly. Two people can walk through the same forest and only one see the plants, insects, and birds and reflect on them. The challenge in mathematics teaching is to ensure that time spent by the students on investigations has the promise of being productive.

Moreover, the general run of schoolchildren, rather than just some elite, has the fundamental right to be involved in being creative in mathematics and genuinely experiencing the fundamental interplay between quasi-empirical and deductive thinking. To achieve this, however, we will need to devote a great deal of care to developing and refining students' mathematical taste and ensuring that they become aware of the different functions of quasi-empirical methods and proof, as well as of the status of the mathematical statements they make and of the strength of the inferences that can be made.

Of course, all of the aforementioned also applies equally to mathematics-teacher education, since the dominant experience of many teachers has unfortunately been that proof is a memorized existing product, not a personal heuristic process. Furthermore, most mathematics teachers have themselves had very little exposure to quasi-empirical work. That is not how undergraduate mathematics courses are usually taught, nor is the balance of proof, conjecture, and investigation a common experience in their preparation for teaching. Indeed, serious thought needs to be given to the development of courses and materials more or less along the lines of the university geometry text by Henderson (1996), which has a really good balance between exploration and convincing argument.

Acknowledgements

A previous, shorter version of this paper, 'The Value of Experimentation in Mathematics,' was published in the proceedings of the Ninth National Congress of AMESA (Association of Mathematics Education of South Africa), 30 June to 4 July 2003, Cape Town, pp. 174–185; available http://mzone.mweb.co.za/residents/profmd/experiment.pdf (accessed April 28, 2004).

I would specifically like to thank the referees, as well as several colleagues and friends, for extremely constructive and valuable comments that assisted me greatly in refining the initial draft. This research was partially funded by a National Research Foundation (NRF, Pretoria) 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).

Notes

1 However, note, on the other hand, that these deductive discoveries by the ancient Greeks illustrate the so-called 'discovery' function of proof, as discussed in De Villiers (1990).

2 In fact, it holds for a tetrahedron with equi-areal faces, and for equi-sided or equi-angled polygons.

3 Of course, sometimes a combination of reflective thought and experimentation is needed. For example, from $3^2 + 4^2 + 5^2$ and $5^2 + 12^2 + 13^2$, we can see that perhaps $7^2 + 24^2 + 25^2$ and guess that similar equations might hold for 9^2, 11^2, 13^2 … Indeed, noting the structure that, say $13^2 - 12^2 = (13 - 12)(13 + 12) = 25 = 5^2$ gives us a clue for constructing and checking individually with a minimum of pain other instances.

4 It would obviously be instructive for students to examine further this conjecture and to identify the additional key property that one of the perpendicular diagonals should bisect the other. This sort of inarticulation of hypotheses happens quite regularly with inexperienced students and requires a fostering of a state of mind characterized by acute analysis and thoroughness.

5 As before, it might be a valuable learning experience to guide students to identify the additional property that the equal diagonals need to cut each other in the same ratio.

6 The problem can also be theoretically analysed more closely to show why the general conjecture is untenable, and it might be instructive to provide students with this additional, different perspective.

7 Though Descartes already, in 1639, knew of the invariance of the so-called 'total angle deficiency' of polyhedra and that the Euler formula can be derived from this, there is, according to Grünbaum & Shephard (1994, p. 122), no real historical evidence that Descartes actually saw the connection.

8 This is a specific case of a Pell equation, for which solutions were discovered as an offshoot of theoretical work rather than quasi-empirical testing. For example, one can see with a modest amount of experimentation that $x^2 - dy^2 = 1$ is solvable in positive integers when d is a small, positive non-square integer and infer (as Indian mathematicians did in the twelfth century) that it is probably solvable for more general d. This led to ad hoc algorithms that worked pretty well (Bhaskara managed the case $d = 61$), and finally, to a theory that produced the present continued-fraction treatment, which is guaranteed to churn out a solution (and will do so with $d = 991$ in fairly short order).

9 Students could be encouraged to explore this problem further to discover (and perhaps prove) that the number of maximum regions is actually given by $1/24(n^4 - 6n^3 + 23n^2 - 18n + 24)$.

10 Unfortunately many people, including many who decide educational policy, have such an impoverished experience of mathematics that they are neither ideologically nor practically disposed to have students engage the discipline at a deeper and more authentic level.

References

Arber, A. (1954). *The mind and the eye.* Cambridge: Cambridge University Press.

Aigner, M., & Ziegler, G.M. (2004). *Proofs from THE BOOK* (3rd ed.). New York: Springer.

Borba, M.C., & Skovsmose, O. (1997). The ideology of certainty in mathematics education. *For the Learning of Mathematics, 17*(3), 17–23.

Burger, W.F., & Shaughnessy, M. (1986). Characterizing the Van Hiele levels of development in geometry. *Journal for Research in Mathematics Education, 17*(1), 31–48.

Bradis, V.M., Minkovskii, V.L., & Kharcheva, A.K. (1959). *Lapses in mathematical reasoning.* Oxford: Pergamon Press.

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

De Jager, C.J. (1990). When should we use pattern? *Pythagoras, 23,* 11–14.

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, UK: Penguin.

De Villiers, M. (1989). From 'TO POLY' to generalized poly-figures and their classification: A learning experience. *International Journal of Mathematics Education in Science and Technology, 20*(4), 585–603.

De Villiers, M. (1990). The role and function of proof in mathematics. *Pythagoras, 24,* 17–24. Available: http://mzone.mweb.co.za/residents/profmd/proofa.pdf (accessed 11 May 2004).

De Villiers, M. (1996). *Some adventures in euclidean geometry.* Durban, South Africa: University of Durban-Westville.

De Villiers, M. (1997). The role of proof in investigative, computer-based geometry: Some personal reflections. In D. Schattschneider & J. King (Eds.), *Geometry turned on!* (pp. 15–24). Washington, DC: MAA.

De Villiers, M. (1998, November). Dual generalizations of Van Aubel's theorem. *Mathematical Gazette,* (496), 405–412.

De Villiers, M. (1999, March). A further generalization of the Fermat-Torricelli point. *Mathematical Gazette, 72*(1), 14–16.

De Villiers, M. (2000, March). A Fibonacci generalization: A Lakatosian example. *Mathematics in College,* 10–29. Available: http://mzone.mweb.co.za/residents/profmd/fibo.pdf (accessed 11 May 2004).

De Villiers, M. (2002). From nested Miquel triangles to Miquel distances. *Mathematical Gazette,* (507), 390–395.

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

De Villiers, M., & Meyer, J. (1995). A generalized dual of Napoleon's theorem and some further extensions. *International Journal of Mathematics Education in Science and Technology, 26*(2), 233–241.

Epstein, D., & Levy, S. (1995). Experimentation and proof in mathematics. *Notices of the AMS, 42*(6), 670–674.

Ernest, P. (1991). *The philosophy of mathematics education.* Basingstoke, UK: Falmer Press.

Eves, H. (1969). The history of geometry. In Hallerberg, A.E. (Ed.), *Thirty-first NCTM Yearbook. Historical topics for the mathematics classroom* (pp. 165–221). Reston, , VA: NCTM.

Fischbein, E. (1982). Intuition and proof. *For the Learning of Mathematics, 3*(2), 9–18, 24.

Fuys, D., Geddes, D., & Tischler, R. (1988). *The Van Hiele model of thinking in geometry among adolescents* (Research Monograph No. 3). Reston, VA: NCTM.

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

Grünbaum, B. (1993). Quadrangles, pentagons, and computers. *Geombinatorics,* (3), 4–9.

Grünbaum, B., & Shephard, G. (1994). A new look at Euler's theorem for polyhedra. *American Mathematical Monthly, 101*(2), 109–128.

Hales, T.C. (2000). Cannonballs and honeycombs. *Notices of the AMS, 47*(4), 440–449.

Halmos, P. (1984). Mathematics as a creative art. In D. Campbell & J. Higgens (Eds.), *Mathematics: People, problems, results* (Vol. 2, pp. 19–29). Belmont, CA: Wadsworth.

Hanna, G. (1983). *Rigorous proof in mathematics education.* Toronto: OISE Press.

Hanna, G. (1989). More than formal proof. *For the Learning of Mathematics, 9*(1), 20–23.

Hanna, G. (1995). Challenges to the importance of proof. *For the Learning of Mathematics, 15*(3), 42–49.

Hanna, G. (1997). The ongoing value of proof. In M. De Villiers & F. Furinghetti (Eds.), *ICME-8 Proceedings of topic group on proof* (pp. 1–14). Centrahil, South Africa: AMESA.

Henderson, D. (1996). *Experiencing geometry on plane and sphere.* Upper Saddle River, NJ: Prentice Hall.

Hilton, P., & Pedersen, J. (1994). Euler's theorem for polyhedra: A topologist and geometer respond. *American Mathematical Monthly, 101*(10), 959–962.

Hofstadter, D.R. (1985). *Metamagical themas: Questing for the essence of mind and pattern.* New York: Basic.
Hölder, O. (1924). *Die Mathematische Methode.* Berlin: Springer.
Horgan, J. (1993, Oct). The death of proof. *Scientific American,* 74–82.
Kleiner, I., & Movshovitz-Hadar, N. (1991). The role of paradoxes in the evolution of mathematics. *American Mathematical Monthly, 101*(10), 963–974.
Kramer, E.E. (1981). *The nature and growth of modern mathematics.* Princeton, NJ: Princeton University Press.
Lakatos, I. (1983). *Proofs and refutations* (4th rpt.). Cambridge, UK: Cambridge University Press.
Lester, J. (1997). Triangles III: Complex triangle functions. *Aequationes Mathematicae, 53*(4), 35.
Luthuli, D. (1996, August). Questions, reflection and problem posing as sources of inquiry in Euclidean geometry. *Pythagoras, 40,* 17–27.
Movshovitz-Hadar, N., & Webb, J. (1998). *One equals zero and other mathematical surprises.* Emeryville, CA: Key Curriculum Press.
Neubrand, M. (1989). Remarks on the acceptance of proofs: The case of some recently tackled major theorems. *For the Learning of Mathematics, 9*(3), 2–6.
Northrop, E.P. (1980). *Riddles in mathematics.* Harmondsworth, UK: Penguin.
Oldknow, A. (1995). Computer aided research into triangle geometry. *The Mathematical Gazette, 79*(485), 263–274.
Oldknow, A. (1996). The Euler-Gergonne-Soddy triangle of a triangle. *The American Mathematical Monthly, 103*(4), 319–329.
Polya, G. (1954). *Mathematics and plausible reasoning: Induction and analogy in mathematics* (Vol. 1). Princeton, NJ: Princeton University Press.
Polya, G. (1968). *Mathematics and plausible reasoning: Patterns of plausible inference* (Vol. 2). Princeton, NJ: Princeton University Press.
Polya, G. (1981). *Mathematical discovery* (Vol. 2). New York: John Wiley.
Rav, Y. (1999). Why do we prove theorems? *Philosophia Mathematica, 7*(3), 5–41.
Rota, G-C. (1997). The phenomenology of mathematical beauty. *Synthese, 111,* 171–182.
Rotman, J. (1998). *Journey into mathematics: An introduction to proofs.* Upper Saddle River, NJ: PrenticeHall.
Schoenfeld, A. (1986). On having and using geometric knowledge. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 225–264). Hillsdale, NJ: Lawrence Erlbaum.
Singh, S. (1998). *Fermat's last theorem.* London: Fourth Estate.
Srinivasan, V.K. (2002). Geese are swans and swans are geese or selecting unripe fruits instead of ripe fruits. *International Journal of Mathematics Education in Science and Technology, 33*(6), 941–949.
Volmink, J. (1990). The nature and role of intuition in mathematics. *Pythagoras, 22,* 6–10.
Wilder, R.L. (1967). The role of intuition. *Science, 156,* 605–610.
Wittmann, E. (1981). The complementary roles of intuitive and reflective thinking in mathematics teaching. *Educational Studies in Mathematics, 12,* 389–397.
Waterhouse, W.C. (1994). A counterexample to Germain. *American Mathematical Monthly, 101*(2), 140–150.