THE WIDTH OF A PROOF

Gila Hanna

Ontario Institute for Studies in Education, University of Toronto, Canada

The paper's aim is to discuss the concept of "width of a proof" put forward by Timothy Gowers. It explains what this concept means and attempts to show how it relates to other concepts discussed in the existing literature on proof and proving. It also explores how the concept of "width of a proof" might be used productively in the mathematics curriculum and how it might fit with the various perspectives on learning to prove.

Keywords: argumentation, explanation, proof

INTRODUCTION

The paper addresses parts of themes one and four of the call for papers, namely "The importance of *explanation*, *justification*, *argumentation* and *proof* and their relationships in mathematics and in mathematics education" and "The use, evolution, elaboration or integration of theoretical frames relevant for research argumentation and proof in an educational perspective". Its aim is to discuss the concept of "width of a proof" put forward by Timothy Gowers (2007).

Both mathematicians and mathematics teachers are well aware that "mathematical explanation" and "mathematical understanding" are elusive terms. Indeed, these concepts are recurring topics in both mathematics and mathematics education. Even though the term "understanding" is not precisely defined, it is often invoked in the context of proof and proving. Most mathematicians and educators do share the view that a proof is most valuable when it leads to understanding (Balacheff, 2010; de Villiers, 2010; Hanna, 2000; Manin, 1998; Mejia-Ramos et al, 2012; Mariotti, 2006; Knipping, 2008; Thurston, 1994). For this reason, mathematicians see proofs less as correct syntactical derivations and more as conceptual entities, consisting of a logical sequence of related mathematical ideas, in which the specific technical derivation approach is secondary.

But when the focus on proof is primarily conceptual, there arises a need for well-defined criteria, apart from the well-known syntactical ones, by which the quality of a proof can be evaluated. In the absence of such well-defined criteria, value judgements are often made on the basis of properties with an aesthetic flavour that have no precise meaning, such as "comprehensible", "ingenious", "explanatory", "elegant", "deep", "beautiful", and "insightful". Clearly, such properties go well beyond the logical correctness necessary to prove a theorem. But in both mathematical practice and mathematics education these quasi-aesthetic properties of proof, ill-defined as they may be, are extremely important, because they speak to the recognition by mathematicians and educators that proofs are “far more than certificates of truth” (Gowers, 2007, p. 37).

Aesthetically pleasing proofs enjoy a privileged position for that reason alone, of course, but Gowers notes that: “…it is remarkable how important a well-developed aesthetic sensibility can be, for purely pragmatic reasons, in mathematical research” (p. 37). The pragmatic reasons include enhanced ease of communication among mathematicians, increased interest, and the very important feature of being memorable. Memorability takes us to the concept of “the width of a proof”.

WHAT IS MEANT BY THE WIDTH OF A PROOF?

In approaching this concept, it is helpful to consider first what drove the mathematician Timothy Gowers to define the quality of mathematical proof for which he coined the term “width”, and the weight he gives to the associated notion of “memorability”. Deploring the fact that there is no well-defined language for expressing value judgments about the desirable “quasi-aesthetic” properties of proof, such as the ones mentioned above, Gowers (2007) set out by trying to clarify what is really meant by the words already in use to describe the quality of a proof. His attempt to come up with “a good theory of informal mathematical evaluation” (p. 39) led him to suggest paying more attention to memory. More specifically, he found it valuable, in formulating a good “theory of mathematical evaluation”, to consider the way mathematicians employ memory in creating and in understanding proofs.

Memorability

In the field of education the terms “memory” and “memorizing” are unfortunately associated with the undesirable notion of rote learning, which by definition eschews both explanation and understanding. But for Gowers memory has an entirely different meaning; in his discourse the term is very strongly associated with explanation and understanding, and he sees it as a crucial tool for mathematical thinking. Gowers notes, first of all, that a memorable proof is “greatly preferable to an unmemorable one” (p. 39), and adds that “memorability does seem to be intimately related to other desirable properties of proof, such as elegance or explanatory power” (p. 40). This raises the interesting question of why some proofs are easier to remember than others, and what role is played in memorability by various other non-syntactical properties such as elegance and ingenuity.

Unlike other properties so often cited, Gowers argues that the concept of “memory” is sufficiently precise to be easily investigated and eventually quantified. First, it is easy to determine whether one remembers how to prove a theorem or doesn’t. Second, it would be feasible for educators or psychologists to find out what features of proofs make them more memorable. Third, memory is clearly related to the background knowledge of the learner (or the mathematician).

Gowers may have gone too far in stating that there is a "very close connection between memorizing a proof and understanding it" (p.40), as if that were firmly established. He states:

> The fact that memory and understanding are closely linked provides some encouragement for the idea that a study of memory could lie at the heart of an explication of the looser kind described earlier. It is not easy to say precisely what it means to understand a proof (as opposed, say, to being able to follow it line-by-line and see that every step is valid), but easier to say what it means to remember one. Although understanding a proof is not the same as being able to remember it easily, it may be that if we have a good theory of what makes a proof memorable, then this will shed enough light on what it is to understand it that the difference between the two will be relatively unimportant. (Gowers, 2007, p.41.)

The concept of memorability leads Gowers to suggest looking at *width*, a term borrowed from theoretical computer science, where it refers to the amount of storage space needed to run an algorithm. (Gowers draws a parallel between "keeping in mind" and "computer storage space", in the belief that storage in computers is analogous to memory in humans.) Thus the "width of a proof" would be a measure of how many distinct pieces of information or "ideas" one has to keep in mind (to hold in your memory) in order to be able to construct or follow a proof, to understand it, and to remember it.

Width

It is important to make the distinction between the width of a proof and its length. Whereas the length of a proof refers to how many lines of argument are needed to prove a theorem, the width refers to the number of distinct ideas one has to keep in mind (or memorize) in achieving that goal. As mentioned above, for Gowers (2007) the term "memorable" is intimately connected with understanding and means "easy to memorize". Thus it becomes important to try to determine which features of a proof might contribute to making it memorable, and in his opinion:

> Some proofs need quite a lot of direct memorization, while others generate themselves …if we think about what it is that makes memorable proofs memorable, then we may find precise properties that some proofs have and others lack. … if mathematicians come to understand better what makes proofs memorable, then they may be more inclined to write out memorable proofs, to the great benefit of mathematics (p. 43).

Gowers suggests paying close attention to how mathematicians might memorize a piece of mathematics in general and a proof in particular. For this reason he goes on to expand on his notion of "memorability" by placing it in the context of mental arithmetic, discussing how one remembers sequences of numbers and how one might go about adding large numbers without writing anything down.

It is sufficient here to mention familiar arithmetic operations in which the "width" of the calculation (that is, the number of pieces of information one must keep in one's memory in order to perform a calculation) can be reduced by appropriate simplifications. For example, the mental multiplication of 47 by 52 could be greatly reduced in "width", according to Gowers, as soon as it is perceived as the difference of two squares. One could represent the product 47 x 52 as the product of (50-3) by (50+3), minus 47, which yields (2500 – 9) - 47. The reduction in "width" occurs because this way of handling the operation means that one has to keep fewer digits in mind than when mentally performing the multiplication 47 x 52 using the customary algorithm. One has eliminated the need to remember multiple digits, at the cost of introducing a single new insightful idea, the concept of the difference of two squares. In Gowers's terms, this new idea has transformed the original large-width operation into a lower-width one.

Examples

Gowers also illustrates his ideas by citing two examples that are more complex and go beyond the realm of numerical calculation, one being the proof of the existence of a prime factorization (a fundamental theorem of arithmetic) and the other being the proof that the square root of 2 is irrational. The latter is shown here.

Proof that the square root of 2 is irrational

There are several proofs of this theorem using different methods, such as proof by infinite descent, by contradiction, by unique factorization, and by the use of geometry. The proof by contradiction is appropriate for mathematics education, and in fact has often been discussed by mathematics educators in the context of the overall notion of proof by contradiction. The following is one version of the proof by contradiction.

> Assume that $\sqrt{2}$ is a rational number. This would mean that there are positive integers p and q with $q \neq 0$ such that $p/q = \sqrt{2}$. We may assume that the fraction p/q is in its lowest terms; it can be written $\sqrt{2} = (2 - \sqrt{2})/(\sqrt{2}-1)$.
>
> Then, substituting $p/q = \sqrt{2}$; $p/q = (2-p/q)/(p/q-1) = (2q-p)/(p-q)$
>
> Because $p/q = \sqrt{2}$, it lies between 1 and 2, we have $q<p<2q$,
>
> It follows that $2q-p < p$ and $p-q<q$
>
> We found a fraction equal to p/q but with smaller numerator and denominator. This is a contradiction, so the assumption that p/q is in lowest term ($\sqrt{2}$ is rational) must be false. ∎

Gowers' point is that this version of the proof contains a step (step 2) that seems to "spring from nowhere", namely the choice to write $p/q = \sqrt{2}$ as the expression "$p/q = (2- \sqrt{2})/(\sqrt{2}-1)$". Clearly this expression is useful, but where does it come from? Clearly it would come from the memory of the trained mathematician. This is a case in which one has to have used one's memory to store an idea for later application. But this proof is nevertheless of lower width than it would have been without this one ingenious idea, because of the greater number of calculations that would have been needed to complete it otherwise.

Of course it would help the reader of the proof to clarify why someone thought of this idea, because a step in a proof should not appear as a *deus ex machina*. Leaving that aside, however, Gowers points out that this particular step is no more than “what mathematicians normally mean by an idea” (p. 48). Ideas, according to Gowers, are in fact an intrinsic property of proofs, side by side with the property of correct logical derivation.

An example from geometry: Proof of Viviani's Theorem

> Viviani’s Theorem: For a point P inside an equilateral triangle ΔABC, the sum of the perpendiculars p_a, p_b, *and* p_c from P to the sides of the triangle is equal to the altitude h.
>
> Proof: The idea is to recall that the area of a triangle is half its base times its height. This result is then simply proved as follows:
>
> $\Delta ABC = \Delta PBC + \Delta PCA + \Delta PAB$
>
> With s the side length, we have:
>
> $\frac{1}{2} sh = \frac{1}{2} sp_a + \frac{1}{2} sp_b + \frac{1}{2} sp_c$
>
> $h = p_a + p_b + p_c$ ∎

This proof of Viviani’s theorem is also of low width, because it can be generated by using just one powerful idea (the area of a triangle).

Back to width

What then is the width of a proof? Gowers’ first try at a definition is:

> It is tempting to define the width of a proof …as the number of steps, or step-generating thoughts, that one has to hold in one’s head at any one time. … suppose that I want to convince somebody else … that a proof is valid, without writing anything down. It is sometimes possible to do this, but by no means always. What makes it possible when it is? Width is certainly important here. For example, some mathematics problems have the interesting property of being very hard, until one is given a hint that suddenly makes them very easy. The solution to such a problem, when fully written out, may be quite long, but if all one actually needs to remember, or to communicate to another person, is the hint, then one can have the sensation of grasping it all at once (p.55).

This restates the idea that the factor that determines the “width of a proof” is the number of distinct pieces of information, or ideas, needed to complete it. In Gowers’ view, the fewer such pieces, the easier it should be to follow a proof or to actively devise one. A proof of low width, with few pieces of information to carry in one’s head, would be superior to one of high width. Often it is a case of introducing a single new but very productive idea that makes it unnecessary to deal with a larger number of less interesting ones.

That does not mean that a proof must be short, because the length of a proof is not identical with its width, as pointed out earlier. Whereas the length of a proof usually refers to the number of deductive steps required to complete it, the width of a proof refers to the number of distinct items, or distinct ideas, that must be kept in memory in order to read, learn and understand the proof.

Reflections on Gowers' notion of width

Gowers's notion of "width" does not seem to capture in a single measure the many non-syntactical properties of proof, such as "comprehensible", "ingenious", "explanatory", "elegant", "deep", "beautiful", and "insightful". Discussion of this point is difficult because none of these terms are well defined, as Gowers points out. The main difficulty, however, seems to be that these terms are far from being synonyms or even overlapping. One could even argue that some of these properties of proof, if not opposites, are at least orthogonal to one another. Very concise proofs are often considered elegant or beautiful, for example, but they are also less likely to be comprehensible or explanatory.

Thus, even if width does prove to be a useful and potentially measurable property of proof, it would not appear to be one that could serve in place of the less measureable non-syntactical properties that mathematicians have always found meaningful and useful. To which of these properties width is most correlated remains an interesting question. An elegant proof, for example, would presumably be considered of low width (and that would be seen as a positive thing). One can imagine, however, a proof designed expressly to be highly explanatory, in which the multiplicity of ideas brought to bear would make it of high width.

The concept of width would also appear to suffer from a lack of clarity as to how different kinds of memorized information – digits, concepts, rules – are to be counted. In the above example of arithmetic multiplication, the point is made that mental use of the normal algorithm requires that many interim digits be kept in mind, implying that the procedure has a large width. The width can be reduced, as Gowers states, by using an ingenious procedure employing the difference of two squares. Now, it is true that in the ingenious procedure there are fewer digits to memorize, but on the other hand one has to have had in mind the notion of the difference of two squares. One could argue that the mental capacity to store such a relatively sophisticated mathematical notion is greater than that needed to store a few digits. Perhaps there is an opening to refine the idea of width by considering whether or not different types of information might be assigned different weights.

Another consideration might be how long a piece of information has to be kept in mind. To take the example of the calculation just discussed, the concept of the difference of two squares must be kept in mind always, so it can be called upon whenever needed, in this calculation or others, while the memory devoted to interim digits can be released as soon as the calculation is finished. The management of computer storage has a strong focus on how long a piece of storage is required. The inspiration for the term "width" came from computer storage, so it might be refined by taking into account the dimension of time.

In considering the concept of width, one is tempted by some examples to say that lower width is good, while others show the value of a richness of ideas, in which case a greater width is better. Be that as it may, mathematics educators already know that it is not enough for students, when they engage in proving, to have mastered the concepts and techniques required to construct valid sequences of logical steps. It is important for them to have a broad grasp of mathematics, so that they can draw upon a reservoir of important mathematical ideas, stored in their memory, and see how to apply them to the mathematical argument they are engaged in constructing.

Because it is a psychological concept, it is not surprising that width is difficult to define. Gowers says as much when he states that a "… precise discussion of width as a *psychological* concept is quite difficult" (p. 51, italics in the source).

An even more basic question that might be raised, however, is whether it is necessary to have quantifiable measures of non-syntactical properties. Mathematicians know an elegant proof or an ingenious proof when they see one, just as mathematics educators are quite capable of deciding which proofs are going to be more useful in conveying to their students important mathematical ideas and their connections.

In commenting on the concept of width, one must keep in mind that Gowers sees his investigation as a work in progress. While conceding that the notion of "width of a proof" is not precise and certainly still lacks a formal definition, Gowers hopes that more work will lead him to come up with a definition "more precise than subjective-sounding concepts such as 'transparent', or even 'easily memorable'" (p. 57).

ASSOCIATED MATHEMATICS EDUCATION RESEARCH

A review of the current literature on proof and proving in mathematics education shows that the concept of the width of a proof is not discussed (understandably so since it is a novel one), but several research papers have dealt, as Gowers does, with the non-syntactical properties of proofs and their importance.

Mejia-Ramos et al *(*2012), recognizing that different proofs of the same theorem improve mathematical comprehension to different degrees, sought to devise an assessment model that could be used to measure the impact of the proof judged more promising from the point of view of comprehension.

They first reviewed the mathematics education research on the purposes of proof; and then the existing recommendations for proof that encourage comprehension, at various educational levels. Finally they interviewed nine university mathematics professors to determine what types of proof comprehension were most valued in university mathematics. Judging by their abstract, these researchers saw the importance to comprehension of the non-syntactical aspects of proof that impelled Gowers to look into the idea of the "width of a proof".

> […]in undergraduate mathematics a proof is not only understood in terms of the meaning, logical status, and logical chaining of its statements but also in terms of the proof's high-level ideas, its main components or modules, the methods it employs, and how it relates to specific examples. We illustrate how each of these types of understanding can be assessed in the context of a proof in number theory. (Mejia-Ramos et al, 2012, p. 3).

Hanna and Barbeau (2009) investigated the properties intrinsic to certain proofs that allow them to convey to students methods and strategies for problem solving. They also looked at the extent to which certain types of proof might yield new insights that are somewhat easier to keep in memory.

Closely related to the concept of width is that of a memorable idea in a proof, which Leron (1983) looked at from a unique perspective in his work on structured proofs. Structured proofs are those that are organized into components in such a way that students can see how each component supports the main thrust of the proof. As Leron saw it, the memorable idea was the entire proof structure, not a step within the proof.

Leron and Zaslavski (2009) discuss the strengths and weaknesses of generic proofs. One of the strengths of generic proofs is that they "enable students to engage with the main ideas of the complete proof in an intuitive and familiar context, temporarily suspending the formidable issues of full generality, formalism and symbolism." (p. 56). Generic proofs contain identifiable (and thus potentially memorisable) main ideas. The discussion in Hemmi (2008) is also closely related to pedagogical properties of proofs and to the way in which students encounter transparency and benefit from it.

In her paper "Key ideas: what are they and how can they help us understand how people view proof?" Raman (2003) characterizes people's views of proof by bringing together two ideas about the production and evaluation of mathematical proof, making a distinction between an essentially *public* and an essentially *private* aspect of proof, and the notion of a *key idea* related to explanatory proofs. She concurs with Gowers that for mathematicians proof is essentially about *key ideas*.

The vast research literature on proof and argumentation also contains many references to the valuable properties of proof that Gowers sought to capture with his idea of width. See the extensive surveys by Mariotti, (2006) and Durand-Guerrier et al., (2012); see also Pedemonte (2007). For example, Douek (2007) discusses the "reference corpus" that is used to back up an argument. This reference corpus might include visual arguments and experimental evidence, and might even depend on a given social context. Douek argues for non-linearity as a model of the mathematical thinking that takes place in the process of proving. This is where one sees a parallel between Douek's view of argumentation and Gowers' description of the quality of a proof, recalling that Gowers pointed out that proofs admit of "artificial" steps that look as if they "spring from nowhere" but turn out to be useful.

Jahnke (2009) argues that it is possible to bridge between argumentation and mathematical proof through explicit discussions of the role of mathematical proof in the empirical sciences, and by allowing students to bring in factual arguments. He says that "giving reasons in everyday situations means frequently to mention only the fact from which some event depends without an explicit deduction" (p. 141).

Knipping (2008) speaks of a method for revealing structures of argumentation in the classroom and invokes the concepts of local and global arguments to analyze a proof, believing that mathematical logic alone cannot capture all the processes of proving. Local arguments can certainly qualify as "proof ideas" in Gowers's sense.

In sum, Gowers' focus on memorability as a desirable property of a mathematical proof, and his introduction of the concept of width, bring a new dimension to the teaching of proof. There remains a need to investigate the relationship between "width of a proof" and memorability on the one hand, and on the other hand the many other important dimensions of proof and proving, such as mastery of mathematical concepts and techniques, grasp of proof structure, and knowledge of methods of proving. A deeper investigation of the concepts of "width of a proof" and memorability promises to lead to new ways of presenting and teaching proofs.

REFERENCES

Balacheff, N. (2010). Bridging knowing and proving in mathematics: A didactical perspective. In G. Hanna, H. N. Jahnke & H. Pulte (Eds.), *Explanation and Proof in Mathematics. Philosophical and Educational* Perspectives (pp. 115- 136). New York: Springer.

De Villiers, M. (2010). Experimentation and proof in mathematics. In G. Hanna, H. N. Jahnke & H. Pulte (Eds.), *Explanation and Proof in Mathematics. Philosophical and Educational* Perspectives (pp. 205-222). New York: Springer.

Douek, N. (2007). Some remarks about argumentation and proof. In P. Boero (Ed.). *Theorems in school* (pp. 163-184). Rotterdam/Taipei: Sense Publishers.

Durand-Guerrier, V., Boero, P., Douek, N., Epp, S. & Tanguay.D. (2012). Argumentation and proof in the classroom. In G. Hanna & M. de Villiers (Eds), *Proof and Proving in Mathematics Education: The 19th ICMI Study* (pp. 349-368). New York: Springer.

Gowers, W. T. (2007). Mathematics, memory and mental arithmetic. In M. Leng, A. Paseau & M. Potter (Eds.) *Mathematical Knowledge* (pp. 33-58). Oxford, GBR: Oxford University Press

Hanna, G. (2000). Proof, explanation and exploration: an overview. *Educational Studies in Mathematics*, 44, 5-23.

Hanna, G., & Barbeau, E. (2009). Proofs as Bearers of Mathematical Knowledge. In G. Hanna & H. N. Jahnke & H. Pulte (Eds.), *Explanation and Proof in Mathematics. Philosophical and Educational Perspectives* (pp. 85-100). New York: Springer.

Hemmi, K. (2008). Students' encounter with proof: the condition of transparency. *ZDM - The International Journal on Mathematics Education*, *40*, 413-426.

Jahnke, H. N. (2009). Proof and the empirical sciences. In F. L. Lin, F. J. Hsieh, G. Hanna, & M. de Villiers (Eds.) *ICMI Study 19: Proof and proving in mathematics education.* Vol. 1, 238-245.Taipei, Taiwan: The Department of Mathematics, National Taiwan Normal University.

Knipping, C. (2008). A method for revealing structures of argumentation in classroom proving processes. ZDM, 40, 3, 427-441.

Leron, U. (1983). Structuring mathematical proofs. *American Mathematical Monthly*, 102, 227-242.

Leron, U. & Zaslavsky, O. (2009). Generic proving: Reflections on scope and method. In F. L. Lin, F. J. Hsieh, G. Hanna, & M. de Villiers (Eds.) *ICMI Study 19: Proof and proving in mathematics education.* Vol. 2, 53-58.Taipei, Taiwan: The Department of Mathematics, National Taiwan Normal University

Manin, Yu. I. (1998). Truth, rigour, and common sense, in H. G. Dales and G. Oliveri (eds.). *Truth in Mathematics* (pp. 147–159).Clarendon Press, Oxford.

Mariotti, A. (2006). Proof and proving in mathematics education. In A. Gutiérrez & P. Boero (Eds.). *Handbook of research on the psychology of mathematics education: Past, present and future* (pp. 173-204). Rotterdam/Taipei: Sense Publishers.

Mejia-Ramos, J.P, Fuller, E., Weber, K., Rhoads, K. & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics*, 79, 1, 3-18.

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed? *Educational Studies in Mathematics*, 66 (1), 23-41.

Raman, M., (2003), "Key ideas: what are they and how can they help us understand how people view proof?" *Educational Studies in Mathematics*, 52, 319-325.

Thurston, W. P. (1994). On proof and progress in mathematics. *Bulletin of the American Mathematical Society*, *30* (2), 161-177.