PROOF: A GAME FOR PEDANTS?

JOANNA MAMONA-DOWNS* & MARTIN DOWNS

**University of Patras, Greece*

In this paper, we propose to examine the types of argument that are deemed acceptable at tertiary level mathematics and under which circumstances, and why the expectancy that a proof is required is sometimes relaxed. We specialise on the status of proof in cases where mathematical modelling takes place, and on tasks whose informal resolution rests on two or more mathematical milieu. On occasion, can the insistence on a proof be regarded as pedantry?

Key words: Proof, Credited Argument, Definitional Tautness, Modelling.

Introduction

Students attending Mathematics courses at the tertiary level often have difficulty to understand the status of proof (Jones, 2000). Further there is evidence that students find the standards applied to proof production somehow arbitrary (Gondek et al, 2009). What is the line between a mere argument and a proof? For this question, some would consider a spectrum might be more appropriate rather than a demarcation line. However, such a viewpoint leaves students in the dark as to when their work should end, or whether their final product qualifies as a proof. It is useful for the researcher to distinguish a proof itself from an argument where basically all the essential ideas behind the proof have been collated but are not articulated in terms of ratified mathematical systems. We shall call such an argument a credited argument; we will discuss our choice of wording in the next section, contrasting it with similar notions employed by other educators. Also there, we will give our position of what a proof is in the context of this paper.

If we acknowledge the notion of a credited argument as well as proof, fairly natural educational issues arise, including the following:

- In what circumstances is credited argument acceptable as a final output? The answer to this question might explain why lecturers can be perceived inconsistent in the level of strictness in argumentation they use.
- We do not say that all proof productions necessarily go through a preliminary stage corresponding to a credited argument. However, for those that do, the backing of the credited argument can give support for the student to formulate the proof, and to appreciate what the proof gives beyond the credited argument.
- One reason that argumentation may end as a credited argument is that translating it into a proof can be deemed not worth the effort. Reasons for this could be that the undertaking would be messy and/or lacking in giving

additional insight. Insisting on the proof in such circumstances might be considered as pedantry. However the lecturer might have good reasons to persist; these reasons, however, may not be immediately appreciated by the students.

- A task may invite a mathematical treatment whereas the context is not strictly ‘mathematical’ itself. Sometimes, the situation can be resolved at the level of a credited argument, where the reasoning retains references to the extra-mathematical context. Because of this, it is not a proof. However, if the task is modelled into another format within a recognized mathematical framework, a proof of the model may be available. Is such a proof, though, to be considered as a proof of the original version of the task?
- The solution of a task may ‘import’ tools from another mathematical perspective from the one that is most natural to assume. The fitting together of different mathematical perspectives sometimes can be made at a perceptual level and as a result aspects of argumentation can be glossed over. In such a case, we have a credited argument but not a proof.

The aims of the paper are rather wide and illustrative in character. The main part of the content consists of descriptions of three tasks that demonstrate the difference between a credited argument and a proof, and the particular issues raised above. In particular, we discuss whether a ‘translation’ from a credited argument into a proof is always merited. Before this, we include a more theoretical section where we state our terms in a more definitional manner and refer to related research.

Background Discussion

This section will define and situate the notion of credited argument that we introduce in this paper. To do this we first consider the notion of ‘truth status’ vis-à-vis proof.

There are many perspectives put forward in the literature concerning how different types of (mathematical) argumentation could be contrasted to proof, or (on the other side) identified with proof. We do not have the space or the expertise to be expansive here. We specialize straight away to consider models of argumentation that are not considered as proofs but in a sense embrace all the central ideas on which a proof would be based. There is a sense of certainty but a lack of usage of the requisite mathematical tools to present a proof. In this way the argumentation obtains a ‘truth status’ but is not completely verified as a proof. Intuition can play a large part in the truth status; Fischbein (1987) says that ‘the concept of intuition …expresses a fundamental, very consistent tendency in the human mind: the quest for certitude’. In the past, an argument that carried conviction both personally and for others was taken as a qualification for proof (Hersh, 1993), or at least conviction is often a prerequisite for seeking a proof (De Villiers, 1990); today, a consensus in terms of conviction could be accepted as an indication of truth but not necessarily for proof (c.f., Segal, 2000). Another notion that educators employ is a warrant. According to Rodd (2000), a warrant is a justification / rationale *for the belief* in mathematical propositions; a warrant acts 'as the lever for (mathematical) knowledge', where 'knowledge generally entails "truth" '. Durand-Guerrier (2008) considers the difference of truth and validity in the context of logical aspects in mathematical proof. In particular, she quotes Tarski: "the truth of a proposition lies in its agreement (or correspondence) with reality; or a proposition is true if it designates an existent state of things".

All the models mentioned above are attributive (rather than operative); they characterize argumentation in terms of a completed output of thought. In this paper, we want to stress the possibility that argumentation promoting confidence in the truth status also can be a basis to form a proof. In this case, care is required both in specifying what argumentation conveys 'truth' and in one's stance of what proof is. We consider proof first. We take a position similar to Thurston (1995) that proof production takes place in the milieu of a certain mathematical language; its' vocabulary is technical but also allows a generous allowance of informal expression in using technical terms. It is difficult for students to attain this language, and 'the language is not alive except to those who use it'. The technical vocabulary is based on explicit mathematical definitions (Section 2.1, Mamona-Downs & Downs, 2005). If an argument in the mathematical language is challenged, then one always has the recourse to argue more rigorously in terms of first - principle definitions and their known logical consequences. We characterize an argument as a proof if it has this recourse; we then say that the argument has 'definitional tautness'.

Now we consider the situation of having a line of reasoning lacking definitional tautness but satisfactorily conveys the truth status; we shall call this a credited argument. This begs the question how we regard the truth status in this context. We find difficulties in confronting this issue directly. The direction we take is to imagine an expert to examine the argument and recognise channels that would allow it to be translated into the mathematical language, as envisaged by Thurston. This stance has unsatisfying aspects; in particular, one would like to say that an individual can feel certainty in his or her argument without exterior authority. However, the reference to a proof clearly is the most reliable source to merit the truth status of an argument. A student, then, might believe that his/her reasoning determines the truth; but this belief must be justified itself, perhaps requiring external ratification. The word 'credited' carries allusions to both 'confidence felt in the veracity in a body' and endorsement, explaining our choice in wording of the term credited argument.

The forming of a credited argument always indicates a lack in the usage of relevant mathematical tools. Sometimes the tools are not available simply because the student has not been introduced to a requisite mathematical theory. This has led some researchers to make a distinction between what is called 'disciplinary proof' (i.e., proof as viewed by professional mathematicians) and 'developmental proof' referring to the stages of learning where different types of tools pertinent to proof making are made. (See ICMI Study 19, discussion document.) The interest of this paper, though, is when the students are familiar with the tools needed; this means that the student potentially is equipped to 'convert' a credited argument into a proof. This suggests a two-stage process in producing a proof, leading to the issues mentioned in the introduction. (There are other models for which two stages are identified in forming a proof; for example, ‘pre-conjecture’ and ‘post-conjecture’ stages are discussed in Pedemonte (2007), but conjectures are not stressed in this paper. Other papers, such as Mamona-Downs & Downs (2009) and Zazkis (2000), consider students’ ability to convert mental argumentation into a proof presentation.)

One of the issues brought up in the introduction was modelling; this might need some clarification. A credited argument might be available within the context (i.e., prior to modelling), whilst the proof only verifies the model (and not necessarily the actual act of modelling). Another term that is often used by educators is ‘mathematization’, that seems to convey the same kind of purpose of rendering a non-mathematical situation into a controlled mathematical environment. However, we regard that an act of mathematization refers more closely to the intuitive reasoning available in the context (compared to an act of modelling), so that it is easier to accept that a credited argument can act as a ‘template’ for the proof. This point will be illustrated later in the paper.

Expositional Examples

We seek to illustrate some issues in the framework of the difference between credited argument and proof, as raised in the introduction.

The first task is widely known, but we want to go further in its solution than usual. The task is:

For a unit square 8×8 array, the bottom-left square and the top-right square are removed. Show that the resultant figure cannot be covered by rectangular tiles all of dimension 1×2.

The usual way to demonstrate this is as follows. Suppose that for the 8×8 array, the squares are painted white or black like a chessboard. Then there are 32 squares painted white, the other 32 squares black. Now the opposite corner squares removed will have the same colour, white say. Hence the resultant figure has 30 squares painted white, 32 black. But each tile must cover one white square and one black. So for a tiling, a necessary condition is that there is an equal number of white tiles as they are black. Hence a tiling cannot exist.

The argument is highly contextual, however carries a strong sense of conviction. What is remarkable is the assertive character it has. Every claim made reads like a fiat; 'it is so'. There could be an objection in that its declarative manner rests much on perception and in this way seems to be at odds to deductive reasoning. At the same time, it seems difficult to get around this. How can you introduce 'definitional tautness' in such a physical situation? This problem seems to be compounded when in the argument itself the notion of colours is brought in; even in the process of solving the problem, another undefined notion is introduced.

Despite this, we believe that many mathematicians would accept the solution as being a valid answer. However, if one were to ask whether it constitutes a proof, the response likely would not be so unanimous. If you gave them time, some probably would say that the problem lies in the task rather than the method: "the task, if you want a fully grounded argumentation, should have been…". Below, we suggest a plausible candidate for such an alternative form of the task:

> Let $S := \{1, 2, \ldots, 8\}$.
>
> Let $T := S^2 \setminus \{(1, 1), (8, 8)\}$.
>
> Let W be the set of subsets of T of two elements of the form:
>
> $\{(u, v), (u + 1, v)\}$ or $\{(u, v), (u, v + 1)\}$.
>
> Prove that there cannot be a subset R of W satisfying both
>
> r_1, r_2 are different elements of $R \Rightarrow r_1 \cap r_2 = \varnothing$ and
>
> $\{t \in T : t \in r \text{ for some element } r \text{ of } R\} = T$.

At first sight, the second task might seem radically different to the original. In a way, indeed it is; where are the objects and actions understood on a physical level that motivated the exercise in the first place? Moreover, if you were presented the second task independently, it would seem highly contrived; which eccentric would dream up such a convoluted seeming creature? No, there is no motivation unless the tasks are regarded as a pair, and it is not so difficult to see how the two are (isomorphically) connected. The sets and the conditions that appear in the second task abstractly are readily 'lifted' to the environment of the first with contextual meaning. Hence, the set T represents the depleted array, W the potential positions in the array that one tile can take, R a tiling of T: the two conditions explain what we mean by a tiling in the set theoretical setting. To prove the proposition, one constructs the sets:

$$T_1 := \{(u, v) \in T : u + v \text{ is odd}\}$$

$$T_2 := \{(u, v) \in T : u + v \text{ is even}\}$$

that accounts for the colouring. By showing that $|T_1| \neq |T_2|$, one proves that R cannot exist. The details here are left to the reader.

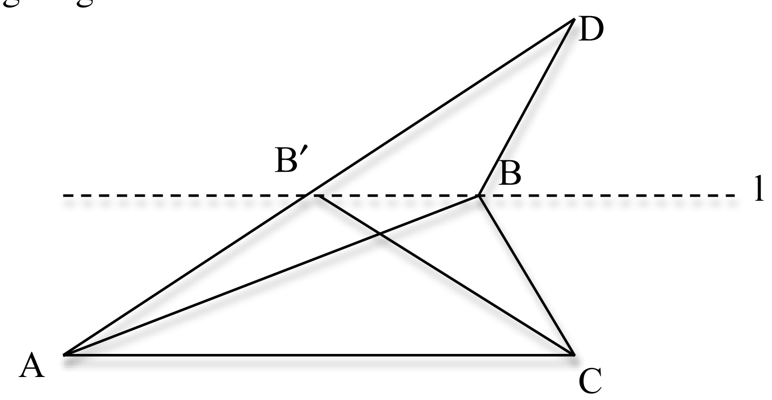
Psychologically, one can either 'identify' the two tasks regarding them essentially the same, or separate the two whilst acknowledging they have completely consonant structure. In the first case, one might say that an act of 'mathematization' has taken place, whilst for the second an act of modelling. The word mathematization suggests that the second version of the task would be regarded as a channel to give the appropriate tools to prove the first task, whereas the modelling viewpoint would suggest that the proof status holds only for the second task. As a central construction transfers in the *solution* (i.e., the notion of colouring to the sets T_1 and T_2) we regard that here mathematization is the more appropriate term to use of the two, as insight is given why the original contextual argument works.

Suppose that a teacher decided to present the original task; after expounding the credited argument, what would be the pedagogical advantages in converting it into a form that permits a proof? Well, the decision whether to stop at the more intuitive and contextual level or to go further of course depends on the teachers' own aims. But the advantages in continuing could include giving students an opportunity to judge what is a proof or not, and impressing them that problem solving is not necessarily restricted to obtain a result but has a role in the forming of mathematization or mathematical modelling. (Obtaining the phrasing of the second task is the result of considerable reflection.) Further, a sense of structural commonality is conveyed, as is the recourse to fundamental mathematical ideas such as a set. Finally, aspects of the modelling or mathematization used in one task could be emulated in other tasks (for example, other problems involving 'arrays' could be susceptible to treatments involving Cartesian products).

Let us now proceed to another example:

Let A, C points of the plane. Let l be any line parallel to AC. Let S be the set of triangles ABC where B is a point on l. Show that the triangle of S that has the least perimeter is isosceles.

This task is conducive to conventional calculus tools for optimization of functions. However we stick to a solution that retains its basis in geometry. We consider the following diagram:



What prompted us to make this diagram? Well, B′ is a point on l such that the triangle AB′C is isosceles. The triangle ABC is drawn as a generic object satisfying the conditions. As the two triangles share the same basis, the problem reduces to show that:

$$2|AB'| \leq |AB| + |BC| \quad \text{for all B on the line l.}$$

To show this, we construct the triangle B′DB that represents the reflection of triangle B′BC in the line l. As a reflection preserves lengths, we have from the resultant triangle ADB:

$$|AD| = 2|AB'| \leq |AB| + |BD| = |AB| + |BC|. \qquad (1)$$

This argument is certainly persuasive, however there are points of awkwardness in it. It rests mostly on Euclidean geometry, for which the recourse to reflection is made mostly on a perceptive level; for instance the fact that the points A, B′ and D are co-linear is assumed or regarded obvious in the argument. The reflection is not treated according to definitional tautness.

A little shift in how to read the constructive elements of the diagram removes these difficulties. Now we suppose that D is the point for which B′ is the mid-point of AD. Consider the triangles B′DB and B′CB. They share lengths of two sides and the angle between, so are congruent. This means that $|BC| = |BD|$ and we are in the position to state the inequality (1) on a firmer ground.

Hence we have demonstrated the proposition via procedures that are in common currency in Euclidean Geometry, and thus we have crafted a proof. We succeeded to circumvent the problem of introducing a reflection. But we have paid a price in doing this. The first argument involving a reflection is more influential in forming the diagram; the second just exploits it.

This example, we contend, illustrates an almost universal phenomenon in formulating a proof; there are 'opposite forces' in basing an argument conceptually against concerns about means of full explication. When ‘foreign’ elements are introduced into a task environment that lends itself to a particular mathematical domain either these elements have to be assimilated, or have to be sidestepped as in the example above. Obviously it is preferable to have a proof that seems transparent once it has been exposed, and this sentiment has been passed on by many mathematicians as well as educators. The phrase " A good proof is one that makes you wiser", accredited to the eminent mathematician Yu. I. Manin, has become almost a maxim. However the situation in reality is more complicated. The advantage in distinguishing a credited argument from a proof is that a teacher or student can gauge which of the two 'makes you wiser', and the acknowledgment that both can be acceptable according to the given circumstance would assuage students' confusion of what constitutes a proof.

An example that shows aspects of students’ behaviour

The example recounted here concerns a student activity. The eight students are, at the time of writing, participants of a master’s program in Didactics of Mathematics; all the students had recently graduated from a Mathematics department. For the activity, a task with a 'realistic' context, an informal argument and a mathematical model is the material given to the students; what was left for the student to do was to produce an argument in terms of the model. (The example serves as an indicator of issues rather than an empirical analysis, so details of methodology are omitted.)

The givens:

Task: Let g(n,k) denotes the number of ways of placing k indistinguishable lions in n cages (in a row)) such that no cage contains more than one lion and no two lions are put in consecutive cages. Show that $g(n,k)=\binom{n-k+1}{k}$ (i.e., the number of ways we can choose k things out n-k+1) .

An informal argument: Suppose that we have a legal positioning of the lions. Then, except possibly for the k^{th} lion, there is an empty cage to the immediate right to each cage containing a lion. Imagine removing these k-1 cages. Then we have n-(k-1) cages remaining and k lions, but now lions can be put into adjacent cages. Hence we are free to place the k lions wherever we want into the n-k+1.

Mathematical modeling of the task: Let S: = {0, 1}. Define $S_{n,k}$ as the subset of the Cartesian product S^n whose elements satisfy exactly k components equal to 1. Suppose $T_{n,k} \subset S_{n,k}$ defined by:

> $(x_1, x_2, \dots ,x_n) \in T_{n,k} \Leftrightarrow \forall\ i \in \{2, 3, \dots ,n-1\}$, if $a_i=1$ then $a_{i-1}=0$ and $a_{i+1}=0$; if $a_1=1$, then $a_2=0$; if $a_n=1$, then $a_{n-1}=0$.

What the students were asked to do was to prove that $|T_{n,k}|$ equals $|S_{n-k+1,\,k}|$.

There are two options for a student to approach this assignment. The first is to translate the sets of the model back to the contextual families of objects and properties (e. g. lions, cages, in a row…) in the original task environment. Then the set theoretical task is treated through the less precise but more semantic setting of its 'isomorphic' realistic-like task. The disadvantage is not only you are shifting from a situation that makes available the tools allowing proof to a situation that is lacking them, but you have the problem to explicitly express the grounds of the transfer itself. The second option is to keep your work within the set theoretical setting, using the contextual version rather like a 'template' to guide the argument but keeping its influence implicit in the exposition. In this case, the direction would be to construct a bijection between that $|T_{n,k}|$ and $|S_{n-k+1,\,k}|$. The advantage here is that you are in a position to give a proof and the argument is direct, a disadvantage (apart from decreased transparency in its presentation) is that what seems immediate in the other task can transfer to messy constructions mathematically.

Of the eight students that participated, seven students clearly took the first option with various degrees of success in expounding the consonant structure between the two tasks. Their reasoning was in terms of the physically understood objects that appear in the description of the original task. Just one student adopted the second option; in his work, clearly the sets were the central actors, with just a few aside references to the context of the first task.

What is the significance of this in educational terms? The authors designed the assignment wishing to test the students' ability to argue in the environment of a model or a system arising from an act of mathematization. The aim was to push forward students' working from a credited argument to a proof via a structurally isomorphic setting. This aim backfired in the way described above. In fact this was anticipated by the teacher (one of the authors); the assignment was set as homework, and the next class was devoted to open debate about the differences between credited argument and proof, and in what circumstances is it useful to try to render a credited argument into a proof. (Such a class debate was possible because it took place within a course, on problem solving and proof, which was part of a master's program in Didactics of Mathematics.) We do not have the space to describe the dialogue that took place, but the overall opinion was that the process of bringing up the mathematical model, or mathematization, was over pedantic. But in forming this opinion, we would have to accept that, in some cases, argumentation without the definitional tautness to qualify it as a proof can give an acceptable mathematical result, even at the tertiary level.

Conclusion

Mathematics students at university are often confused about the nature of proof, and worry whether what they write for a solution is in an acceptable form from the point of view of their teacher. Such confusion is natural because teachers can be inconsistent in what they accept. (For example, the notion of diagrammatic proof is accepted by some mathematicians, not by others.) Subtle hints of the level of deductive reasoning expected can be conveyed; for example, if the directive of a task is in the form 'show that', rather than 'prove that', there is an expectancy that a more relaxed argumentation is allowed. There is a lot of 'etiquette of standards' that students are supposed to pick up by themselves: this clearly pertains to the notion of didactical contract, due to Brousseau (1984). In reality, perhaps mathematicians on their own are not equipped to convey to their students what standards are demanded for different circumstances; educators should take the initiative to assist.

In this paper, we aimed to explain why for certain types of tasks an argument naturally ends even when its form falls far below the level expected for a proof. Here we did not insist on formal proof, which we regard as an ideal that in practice is rarely respected. However, we did insist that a proof is put on a firm mathematical basis where all objects and actions concerned can be explicitly defined. We pointed out and illustrated some situations where the need for a proof is debatable. In particular, we considered problems that are not trivial at all but are conducive to mental argumentation embracing elements that are perceptual in character; also we considered cases where the most natural approach involves more than one mathematical tradition or theory. Are we obliged to model in the first case, and alter the argument such that it fits within a consistent mathematical setting in the latter? We leave this question as an open issue, but we stress that, for students, there is a danger of seeming to be engaged in a game that only a pedant would be interested in. This problem is aggravated in the case where we model one task by another; are we proving the model or the original? In this respect, we raised the issue of the relative meanings of the terms 'mathematical modelling' and 'mathematization' that deserves more research inquiry.

References

Brousseau, G. (1984). The crucial role of the didactical contract in the analysis and construction of situations in teaching and learning mathematics. In Steiner, H. (ed.), *Theory of Mathematics Education, Paper 54*. Institut fur Didaktik der Mathematik der Universitat Bielefeld, pp. 110-19.

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

Durand-Guerrier, V. (2008). Truth versus validity in mathematical proof, *ZDM The International Journal on Mathematics Education* 40 (3), 373-384.

Fischbein, E. (1987). *Intuition in Science and Mathematics*. Dordrecht, Holland: Kluwer Academic Publishers, p.13.

Gondek, V. & Rösken, B. & Törner, G. (2009). Algebra vs. Analysis: University Students' Views on Role of Proofs. *Proceedings of the ICMI Study 19: Proof and Proving in Mathematics Education*. Volume I, pp. 178-183.

Hanna, G. et al. (2009). ICMI Study 19: Proof and Proving in Mathematics Education: Discussion Document.

Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics*, 24 (4), 389-399.

Jones, K. (2000). The student experience of mathematical proof at university level. *IJMEST*, 31(1), 53-60.

Mamona-Downs, J. and Downs, M. (2005). The identity of problem solving. *Journal of Mathematical Behavior 24*, 385-401.

Mamona-Downs, J. & Downs, M. (2009). "Necessary Realignments from Mental Argumentation to Proof presentation". *Proceedings of CERME 6, electronic form*.

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

Rodd, M. M. (2000). On Mathematical Warrants: Proof does not always warrant, and a warrant may be other than a proof. *Mathematical Thinking and Learning*, 2(3), 221-244.

Rota, G. C. (1997). *Indiscrete Thoughts*. Birkhauser, Boston, p. 142.

Segal, J. (2000). Learning about Mathematical Proof: Conviction and Validity. *Journal of Mathematical Behavior*, 18(2), pp. 191-210.

Thurston, W.P. (1995). On Proof and Progress in Mathematics. *For the Learning of Mathematics*, 15 (1), 29-37.

Zazkis, R. (2000): Using code-switching as a tool for learning mathematical language. *For the Learning of Mathematics,* 20 (3), 38-43.