

“IS THAT A PROOF?”: AN EMERGING EXPLANATION FOR WHY STUDENTS DON’T KNOW THEY (JUST ABOUT) HAVE ONE[†]

Manya Raman^{*}, Jim Sandefur^{**}, Geoffrey Birky^{**}, Connie Campbell^{***}, Kay Somers^{****}

Umeå University^{*}, Georgetown University^{**}, Millsaps College^{***}, Moravian College^{****}

This paper describes an episode taken from the third year of a design experiment aimed at improving the teaching and learning of proof at the university level. In the episode, students come enticingly close to having a proof, at least as judged by competent outsiders. However the students themselves, while satisfied with their result, abandon it when asked to write up a formal proof. We offer an analysis of this episode and offer questions for further study.

Key words: Proof, Tertiary Level, Key Ideas, Technical Handle, Design Experiment

INTRODUCTION

Design experiments, or “developmental research” as this work is often called in Europe, are becoming increasingly common, at elementary, secondary, and even tertiary education (e.g. Brown 1992, Collins 1999, van den Akker, Branch, Gustafson, Nieveen, & Plomp, 1999, Lesh 2002). The goal is to find theoretically grounded answers to practical questions of the classroom, done in as natural a setting as possible, with as Brown puts it, the “the blooming, buzzing confusion” that one can sometimes find in real classrooms, under real pressures, with real constraints and opportunities.

While the potential of merging theory and practice is quite alluring for many reasons, the practical and conceptual realities of doing so remain challenging. As Kelly (2002) suggests: if design experiments began in the early 1990’s as a sort of art, they are emerging in recent years as a type of science, guided by increasingly rigorous methodology and increasingly useful results. But specifying exactly what this science consists in, that is, how to merge research and practice in a mutually advantageous way, is still a matter of debate, discussion, and development.

This paper is an emerging product of a design experiment aimed at improving the teaching of proof at the university level. The research team, consisting of two

[†] This research is supported by NSF grant number DUE-0736762.

mathematics educators and three mathematicians, came together with the aim of improving the teaching of “Introduction to Proof” courses, a type of course used frequently in American universities to help students prepare for the rigor of the theoretical courses like abstract algebra and analysis¹. The idea was to use videos of students struggling, and eventually succeeding, at proving claims that are known to be hard for students in this type of course, as a basis for discussion. These videos can be used both as a professional development tool for teachers who want to better understand student difficulties with proof and as a curriculum resource for class discussion to help students be more aware about their own mathematical thinking.

After three rounds of testing and piloting, we now have a fairly stable set of curricular materials, which include (1) carefully edited videos of students working on proofs that many other students find difficult, (2) materials to help teachers use these videos, both for their own understanding of student thinking and for classroom use. These materials have been tested in four colleges in the United States in the context of “Introduction to Proof” courses taught by members of the research team and their colleagues. We also are generating a number of research articles, this being one example, that probe questions of mathematical thinking that enhance and/or inhibit proof production.

We consider our particular marriage of theory and practice to be a happy one. The central questions which drive our research—how to reconcile student and faculty thinking about proof and proving—grew naturally from our experiences as teachers struggling to make the best of our own “Introduction to Proof” courses. While not eliminating common sense and experience as legitimate grounds for interpreting data, we felt a real need to move into theoretical territory to help make sense of some of the mysteries of mathematical thinking.

This paper describes one part of this theoretical journey. We begin by describing an episode that our team found particularly compelling. In the episode, students come enticingly close to finding a proof but do not seem to notice that they have done so. Rather than convert what outside observers recognize as a “key idea” of a proof into a formal proof, they abandon the idea and take a different, and ultimately unsuccessful path. This episode is useful for a starting point in understanding the nature of key idea in the process of proof production, but also points to some fuzziness about the notion of key idea, which a more theoretical analysis can help clarify.

FRAMING

The methodology for this project, for which this paper is one small part, follows the program set out by Cobb, et al (2003). The design is highly interactive and

¹ See Alcock (2007) for a similar project, focusing primarily on professional development, which has been successfully piloted in the US and UK.

interventionist, involving gathering and indexing of longitudinal data from a number of sources, including videos of classroom practice, individual and group interviews with teachers and students, journal and email records from the teachers, written records of student work, and audio and video records of behind-the-scenes discussion among the research team. Like Cobb, et al, we see this design experiment as a “crucible for the generation and testing of theory.” It is the tangible pressures of classroom realities that provide a needed spark for the theory to develop and crystallize, and one of the goals of this paper is to make part of that process visible to both research and practitioner communities.

The central research questions involve characterizing the trajectory of proof development in a way that both helps us see where students sometimes go wrong and also gives some guidance towards how to teach students to prove in a more effective way. In particular, as we traced one particular episode in which students struggled, came close, and eventually failed to find a proof we wanted to know (1) what were the critical “moments” when there was opportunity for the proof to move forward, and (2) what is the nature of these moments. In the end we found three such moments, which seem to play a critical role in proof production. These moments do not necessarily occur in every proof, nor do they necessarily occur in the order in which we present them, but they seem to be critical in the sense that if one is present, the proof can move forward in a fairly significant way, and if one is absent, it is quite possible that the proof will not move forward (or that a proof will be produced without a full sense of understanding).

The first moment is the getting of a key idea, an idea that gives a sense of “now I believe it”². The key idea is actually a property of the proof, but psychologically it appears as a property of an individual (we say that a particular person “has a key idea” if it appears that they grasp the key idea of a proof.) We refer to “a” key idea rather than “the” key idea, because it appears that some proofs have more than one key idea. While a key idea engenders a sense of understanding, it does not always provide a clue about how to write up a formal proof.

The second moment, is the discovery of some sort of technical handle, and gives a sense of “now I can prove it,” that is, some way to render the ideas behind a proof communicable³. The technical handle is sometimes used to communicate a particular

² More elaborated discussions of “key idea” can be found in Raman (2003), Raman & Weber (2006), and Raman & Zandieh (in progress). A key idea can be thought of as a certain kind of intuition that has both a public and private character: public in the sense it can be mapped to a formal proof, private in the sense that it is personally understandable as a sort of primary, or *prima facie*, experience. For a careful discussion about intuitions see Bealer (1992).

³ The term “technical handle” here is akin to the term “key insight” in Raman & Weber (2006). We have chosen to change the term in part because it sounded too similar to “key idea” which has a very different character, and in part because the technical aspect of this “moment” seemed central to

key idea, but it may be based on a different key idea than the one that gives an ‘aha’-feeling, or even on some sort of unformed thoughts or intuition (the feeling of ‘stumbling upon’.)

The third moment is a culmination of the argument into a standard form, which is a correct proof written with a level of rigor appropriate for the given audience. This task involves, in some sense, logically connecting given information to the conclusion. We assume that for mathematicians the conclusion is probably in mind for most of the proving process. But for students, the theorem might sometimes be lost from sight, adding a sense of confusion to their thinking processes.

In the data below we will illustrate how each of these moments occurs in the midst of proof production before turning more critically to trying to understand the nature of key idea.

THE EPISODE

The following example illustrates the presence and/or absence of these three moments as students work on the following task:

Let n be an integer. Prove that if $n \geq 3$ then $n^3 > (n+1)^2$.

Students were videotaped working on this task in the presence of the research team, and upon their completion, were asked questions about their thinking. Afterwards, the research team watched and discussed the videos. We were drawn to one part of the proof process that turned out to be a genuine mystery—an episode, near the beginning, in which the students generate what the faculty identify as a correct proof, but what the students, at least at some level, do not recognize as one.

Details: In the first two minutes of working on this task, the students made an observation that the professors identified as a key idea of the proof, namely that a cubic function grows faster than a quadratic. Rather than trying to formalize this idea, the students switched to an algebraic approach, what we label as a technical handle, to try to get to a proof. They wrote $n^3 > n^2 + 2n + 1$ which they manipulated into $n(n^2 - n - 2) > 1$ and then $(n-2)(n+1) > 1/n$.

The students then noticed that if $n \geq 3$ then the terms on the left are both positive integers so the product is a positive integer. And since n is an integer greater than 2, the right hand side is going to be between 0 and 1. They wrote these observations as

if $n \geq 3$ (line break) $n-2 > 0$ (line break) $n+1 > 0$ (line break) $0 \leq 1/n \leq 1$

its nature. The distinction between “key idea” and “technical handle” might appear at first sight to be similar to the distinction Steiner (1978) makes and Hanna (1989) builds on between proofs that demonstrate and proofs that explain. However, it is possible that a key idea gives rise to a proof that demonstrates or explains, and a technical handle can also lead to both kinds of proofs.

and seemed quite pleased with their reasoning, one student nodding and smiling as the other one wrote the last line.

S2: Yeah.

S1: This is if n is greater than 3, if n is greater than or equal to 3.

S2: Yeah.... Cool.

At this point in the live proof-writing, the three professors were convinced that the students had a proof. They believed that “all” the students needed was a reordering of their argument. To show $n^3 > n^2 + 2n + 1$, it suffices to show $(n-2)(n+1) > 1/n$, which one can establish by showing that the left-hand side is a positive integer while the right is between 0 and 1.

However, it turned out that the students, despite being pleased with their argument, were less than sure that they were near a formal proof. A professor asked the students “Is that a proof?” and S1 replied, “That’s what I’m trying to figure out.” As the students moved to now write up the proof, they switched to a new track, trying a proof by contraposition. This attempt ended up turning into a confusing case analysis in which they tried to prove the converse of the contrapositive and investigated many irrelevant cases.

AN EVOLVING EXPLANATION

That students can come so close to a proof without recognizing it is probably familiar to most experienced teachers⁴. *Why* the students are not able to recognize that they are so close is another, more difficult, question. Here we show how looking at the three “moments” of the proof, described above, allows us to compare what the students did in this problem with an idealized version of what faculty might have done.

The moments are represented graphically in Figure 1 below, with the blue line representing the “ideal” (professor-like) proving process, and the red line representing the students’ process⁵. The marks m_i indicate the points in the proof at

⁴ Another example can be found in Schoenfeld (1985) where two geometry students have what the researcher is convinced is a correct “proof” but when asked to write it up, they draw two columns and abandon all their previous work.

⁵ In creating this “idealized” version of a proof, we depict a continuity between the key idea and the technical handle, although we realize in practice that many proofs are made without the author being able to connect the two. The question about whether there exists such a connection, even if it has not been found, is an open one. We also realize that the process of proof development is not linear, even for an able mathematician, in many cases. This picture points out more the over-all trajectory of the proof, with minor false-paths ruled out. Further the heights of the peaks could vary.

which different moments are achieved: m1 for the key idea, which both faculty and students achieved (though the students may not realize this), m2 for the technical handle (which students in this case see as disconnected from their key idea), and m3 for the organization of the key idea and/or technical handle into a clear, deductive argument (which in this case the students never reach.)

Specifically, m1 is recognizing that cubic functions grow faster than quadratic ones. m2 is choosing an algebraic approach, factoring the polynomials before and after the inequality sign. We label this as a technical handle even though the students do not know from the beginning where this might lead⁶. m3 is connecting the assumption that $n \geq 3$ with the conclusion that $n^3 > (n+1)^2$. In this case, the students never reached m3, and in fact—during their attempt to write a formally accepted proof, they seem to lose sight of what they are proving.

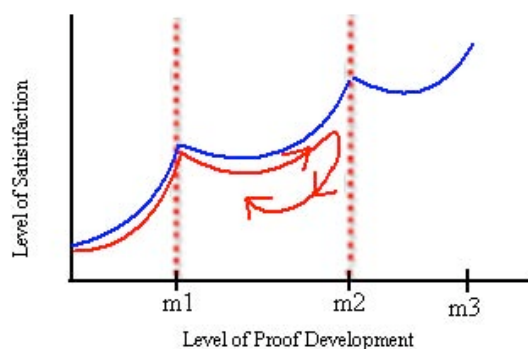


Figure 1: Comparing student (red) and faculty (blue) proof strategies

In the episode above, the students find two key ideas: one that cubics grow faster than quadratics, and another, after students have written $(n-2)(n+1) > 1/n$, that the right-hand term is trapped between 0 and 1 while the left grows indefinitely. Neither of these ideas gets developed into a formal proof. The curved line between m1 and m2 represents how students move towards a technical handle and end up at the second key idea.

⁶ The labeling of technical handle here is a bit tricky. If the students are not themselves aware of the way to link their algebraic manipulation to a proof, is it misleading to say they have found a technical handle since technically they do not seem to register that they “know” how to prove it. We have tentatively labeled this moment as a technical handle anyway, in part because as outside observers we can see that this algebraic manipulation could lead to a correct proof. In addition, while the students might not see exactly how to extract a formal proof from their algebraic arguments, they seem to take their arguments to be convincing and that they have grounds for making a formal argument.

The crucial distinctions between the “ideal” graph and the “student” graph are the breaks at m1 (students do not try to connect their key idea to a technical handle) and m2 (students lose sight of the conclusion and end up trying to prove a converse.) Our data indicate that these breaks are not merely cognitive—it isn’t that the students do not have the mathematical knowledge to write a proof, since they articulate the essence of the proof after three minutes. The problem is epistemological—they don’t seem to understand the geography of the terrain. Expecting discontinuity between a more intuitive argument and a more formal one, the students abandon their near-perfect proof for something that appears to them more acceptable as a formal proof.

Of course it is not always possible to connect key ideas to a technical handle, or to render a technical handle into a complete proof. But what distinguishes the faculty from the students is that the faculty are aware that this connection is possible, and might even be preferable given that sometimes it takes little work—in this case a simple reordering of the algebraic argument would suffice for a proof. As one professor in the study said:

“It became clear that to formalize meant something different to them and to us. To us, formalize seemed to mean ‘simply clean up the details’. To them, it seemed to mean ‘consider rules of logic and consciously use one’.”

Recognizing the difference between radical jumps that need to be made to move mathematical thinking forward and local jumps that allow one to delicately transform almost rigorous arguments into rigorous ones might be an essential difference that mathematics teachers can learn to recognize, diagnose, and communicate to their students.

FURTHER QUESTIONS

The episode and analysis described above, raise a number of questions which we would like to discuss briefly here.

1. Nature of key idea/technical handle

One nice feature of the episode above is that the identification of key idea and technical handle came fairly easy, with relatively little debate or discord among members of the research team. But are the notions of key idea and technical handle so clear that they can be picked out in any setting, for any proof? For this we need to continually refine the definitions (and in this paper we have actually backed away from a technical definition and given more general descriptions.) An ongoing research project of our team involves looking at a broad number of theorems, identifying key ideas and technical handles for different proofs, and refining the definitions based on that data.

2. Context of discovery vs. context of justification

The distinction between context of discovery and context of justification⁷, which has had a significant influence on epistemology and related fields, might be useful for understanding why students do not realize they have a proof. Taking the distinction to be psychological (which was not the original intent, but serves our purpose here), it seems natural to suggest that in the process of proving one has a phase of discovery and a phase of justification.

In the episode above, the students seem to be missing an important half of this combination. They sort of “discover” the key idea without seeing it as a justification⁸. Perhaps being able to toggle between the different contexts is a marker for mathematical maturity, and somehow central for being able to identify a proof as a proof. Specifically, the key idea might involve some combination of seeing the idea as a product of discovery and a grounds for justification (a thing to be justified). This is just a hypothesis, and a more careful analysis of the distinction between discovery/justification is needed to be able to substantiate it.

3. A Fregean telescope?

Another way of seeing the difference between student and faculty understandings in this episode might have to do with a deep connection (or lack of connection) between mathematical objects and the way they are grasped by the mind. This suggestion is highly tentative: to use Frege’s distinction between “sinn” (roughly, sense) and “bedeutung⁹” (roughly, reference) to better understand this relationship (Frege (1892/1997)).

Frege uses the following analogy to explain the difference between sinn and bedeutung: imagine a person looking at the moon through a telescope. The moon is a bedeutung, an object in the world, with a public status. The image on our retina is a sinn, the personal sense we have of that object, which has a private status. The telescope is sort of like a thought that connects the two—it has public status, in the sense that anyone can look through it, but it somehow makes an otherwise difficult to grasp object intelligible to the human mind.

Without going deeply into the way Frege extends this analogy to mathematics (in part because there are tricky moves, both going from the bedeutung of an *object* to the bedeutung of a *sentence*, and going from natural language to

⁷ For the original distinction see Reichenbach (1938), and for a critical discussion of this distinction in contemporary philosophy and history of science see Schickore & Steinle (2006).

⁸ Wright (2001) warns about misinterpreting the word “discover”. He points out that we would not say someone “discovered” the South Pole if they did not realize it was there. It is with this warning in mind that we use the term “discover” in quotation marks.

⁹ We retain the German names since the English translations are not completely accurate.

mathematical language) it might be useful to think if there is an analogy to the telescope in the episode described above.

Could it be that students stand facing some (to them) far away star, and with the aid of a telescope the public could be rendered private? If so, what would the telescope be, and is it something that we could better encourage students to develop and/or use as they learn to prove? Or is it possible that there is no telescope at all, just as when I look at the coffee mug on my desk, I feel I am simply getting sense data of the mug, without any mitigation. Perhaps the mind simply grasps key ideas. If so, then, what explains why some people grasp them and others don't?

This is perhaps merely a rephrasing of the central mystery found in the episode above. But by placing this mystery in a Fregean context (which also allows access to his critics), perhaps we gain some conceptual tools to try to better understand, not only the mystery, but also what we can do about it.

These questions mark a few of the places where we think it might be productive to push for a deeper analysis and where we see possibilities to connect to existing research. We are especially excited about the potential to use results from the field of epistemology where questions about the relation between mental representations and the external world (of which we consider mathematics to be a part) have been discussed extensively. In the next phase of our project, we plan to devote increasing time to developing and refining our theoretical ideas. We welcome any and all suggestions that can help us do so.

REFERENCES

- Alcock, L. (2007). How do your students think about proof? A DVD resource for mathematicians. In J.-H. Woo, H.-C. Lew, K.-S. Park & D.-Y. Seo (Eds.), *Proceedings of the 31st Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 1-8). Seoul, Korea: PME.
- Bealer, G, Strawson P.F. (1992). The Incoherence of Empiricism. *Proceedings of the Aristotelian Society, Supplementary Volumes*, (Vol. 66, pp. 99-143).
- Brown, A. (1992). Design Experiments: Theoretical and Methodological Challenges in Creating Interventions in Classroom Settings. *Journal of the Learning Sciences*, 2 (2), 141–178.
- Cobb, P., Confrey, J., DiSessa, A., Lehrer, R. & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, 32 (1), 9–13.
- Collins, A.(1999). The changing infrastructure of education research. In E.C. Lageman & L.S. Shulman (Eds.), *Issues in education research: Problems and possibilities*. SanFrancisco: Jossey-Bass.

- Frege, G. (1892/1997). On Sinn and Bedeutung, in Beany Michael (Ed.), *The Frege Reader*, Blackwell Publishers.
- Hanna, G. (1989). Proofs that prove and proofs that explain. In G. Vergnaud, J. Rogalski & M. Artigue (Eds.), *Proceedings of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 45-51). Paris: PME.
- Kelly, A. (2004). Design Research in Education: Yes, but is it Methodological? *Journal of the Learning Sciences*, 13 (1) 115-128.
- Lesh, R. (2002). Research Design in Mathematics Education: Focusing on Design Experiments in Handbook on International Research in Mathematics Education, Lyn English (Ed.) Mahwah, NJ: Lawrence Earlbaum.
- 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.
- Raman, M. & Weber, K. (2006). Key ideas and insights in the context of three high school geometry proofs. *Mathematics Teacher*, 99 (9), 644.
- Raman, M. & Zandieh, M. (in progress). The Case of Brandon: The Dual Nature of Key Ideas in the Classroom. Presented at PME-NA Toronto, Oct. 21-24, 2004.
- Reichenbach, H. (1938). Experience and Prediction. Chicago: Chicago University Press.
- Schickore, J. & Steinle, F. (2006). *Revisiting Discovery and Justification: Historical and Philosophical Perspectives on the Context Distinction*. Dordrecht, Netherlands: Springer.
- Schoenfeld, A. H. (1985). *Mathematical problem solving*. Orlando, FL: Academic Press.
- Steiner, M. (1978). Mathematical Explanation. *Philosophical Studies*, 34, 135-151.
- van den Akker, J.J.H., Branch, .M., Gustafson, K., Nieveen, N., & Plomp, T. (Eds.), (1999). *Design approaches and tools in education and training*. Dordrecht, Netherlands: Kluwer.
- Wright, L. (2001). Justification, Discovery, Reason, and Argument. *Argumentation* 15, 97-104.