UNIVERSITY OF CAMBRIDGE
Faculty of Education

ISSN 2043-8338

Journal of Trainee Teacher Education Research

A case study exploring students' conceptions of proof within a Year 10 Mathematics classroom

Darion Mayes

(PGCE Mathematics, 2022-2023)

email: darionjmayes@cantab.net

Abstract

The difficulty in the transition to studying mathematics at university has been documented and studied by a number of previous authors. In this paper, one aspect of this transition – the study of mathematical proof – is examined in more depth through a case study of a high attaining Year 10 class. The results of this case study, whilst necessarily tentative, indicate that students exhibit a pattern of misconceptions regarding proof which is corroborated by larger-scale studies both nationally and internationally. Furthermore, data collected from student interviews during this case study is used to suggest a novel explanation of some of these misconceptions. This explanation connects mathematical proof to broader topics, such as maths anxiety, which potentially merits further investigation.

https://doi.org/10.17863/CAM.109346

A case study exploring students' conceptions of proof within a Year 10 Mathematics classroom

Darion Mayes

Introduction

Within mathematics, proof has a special place – and a special meaning. In this case study, I define a *mathematical* proof as an activity wherein a mathematical result is obtained via a sequence of logical inferences from previously accepted theorems and axioms. Dawkins and Weber (2017, p.132) trace this idea back to ancient Greece, arguing that "Euclid's *Elements* has always served as the paradigm" for deductive proof. As a more modern influence, they cite the philosopher Immanuel Kant for arguing that "only mathematics could produce new knowledge that was independent of experience, providing mathematical knowledge with a more secure footing" (Dawkins & Weber, 2017, p.128).

More recently, the concept of mathematical proof has become embroiled in a series of debates regarding transitions to studying mathematics at university. For example, Solomon (2006, p.377) argues that it is "the focus of a number of well documented complaints regarding students' difficulties in encountering degree-level mathematics". Dawkins and Weber (2017, p.133) concur, adding that "there is ample evidence that students at all levels have difficulty producing these classroom proofs". Earlier such debates led to a publication by the London Mathematical Society (1995, pp.9–10) in which they argued that "most students entering higher education no longer understand that mathematics is a precise discipline in which exact, reliable calculation, logical exposition and proof play essential roles".

The extent to which these criticisms have been addressed by the National Curriculum is unclear. On one hand, an explicit aim of the curriculum is to ensure that all students can "reason mathematically by following a line of enquiry, conjecturing relationships and generalisations, and developing an argument, justification or proof" (Department for Education/DfE, 2013, p.3). On the other hand, the construction of proofs is non-statutory (DfE, 2013, p.5), meaning that the extent to which proofs are taught within the school is heavily dependent on the teachers. Matters are further complicated by the fact that the National Curriculum itself only runs up to GCSE, whereas A Level Mathematics (and Further Mathematics) are more immediate precursors to university-level mathematics.

In this paper, I will investigate the role of mathematical proof within a high-attaining Year 10 Mathematics classroom. In particular, I will explore the ways in which these students conceptualise and interact with proof, and to what extent they diverge from those of more advanced mathematicians. Through this analysis, I aim to identify specific misconceptions of proof within the classroom, which may be used to inform future teaching and professional development.

The remainder of this paper is structured as follows: Firstly, I will conduct a review of previous research surrounding students' conceptions of proof. This will be followed by a discussion of the methodology of the case study, before detailing its findings. Finally, I will discuss the ramifications of these findings, and their implications for my future teaching.

Literature Review

Investigations into mathematical proof form a deep and rich vein of educational research. The aim of this section is to analyse some of the important elements of this literature. In particular, this paper draws upon the investigations of Harel and Sowder (1998) and Healy and Hoyles (2000) on students' conceptions of proof.

Proof Schemes

Harel and Sowder (1998, p.237) argue that one reason for the perceived deficit in students' proof skills is that "we, their teachers, take for granted what constitutes evidence in their eyes". Consequently, they were interested in mapping the different ways (or cognitive schemes) in which students understood mathematical proof. This was done through a series of teaching experiments conducted by the first researcher during courses delivered to mathematics undergraduates in the United States.

The findings of this research are described by the authors as "exploratory", who stress that it "must be validated by other researchers", whilst noting that it has been "cross-checked through interviews of mathematics majors at a separate institution" (Harel & Sowder, 1998, p.238). This reflects an *exploratory sequential design* to the research, characterised by Cohen et al. (2017, p.39) as a research methodology in which "qualitative data are usually collected first (usually with a small sample), with quantitative data from a larger sample used to generalise the findings".

It will be useful to filter criticisms through the lens of this design philosophy. For example, one criticism of the research is that the focus on mathematics undergraduates – rather than high school students – raises questions of generalisability. For the authors, this choice has fewer ethical complications, which potentially simplifies the collection of qualitative data. Moreover, the age of students did not appear to be a particular concern; the authors conclude that proof schemes are a result of "years of instruction" and that instructional activities intended to modify them "must begin in an early age" (Harel & Sowder, 1998, p.280). Nonetheless, I will corroborate those proof schemes which correspond to *misconceptions* of proof using research conducted on younger students.

Similarly, the sample size of the research does not lend itself to statistical analysis, as only one of the teaching experiments meets the minimum of 30 cases suggested by Cohen et al. (2017, p.203). When the experiments are viewed collectively, the sample size is more encouraging, but there is an additional nuance – it is unclear how many students were involved in multiple teaching experiments. For the authors, this may be less problematic, as the data may be used to "document the progress of students" in order to "offer principles for instructional treatments that facilitate proof understanding, production, and appreciation" – two of their stated secondary aims (Harel & Sowder, 1998, p.237).

Conversely, one strength of their research is the breadth of data collected, including video-taped classroom sessions and interviews, written homework and tests, and observation notes by both the researcher and graduate students (ibid. p.238). This variety of methodologies allows for greater triangulation, which mitigates some of the risks inherent to qualitative data. For example, obtaining observational data from multiple independent observers reduces the risk of bias, thus addressing one of the main weaknesses of observations as articulated by Cohen et al. (2017, p.544).

Based upon their teaching experiments, Harel and Sowder (1998, p.245) identified seven different proof schemes. In the interests of time, this paper will focus on three of these proof schemes, each of which correspond to particular misconceptions of proof which appear frequently throughout other literature.

The first such proof scheme is the *authoritarian* proof scheme, wherein a student's "main source of conviction" of a mathematical result is "a statement appearing in a textbook or being uttered by a teacher" (ibid., p.247). Furthermore, they suggested that this proof scheme manifested in a variety of ways, such as an aversion to trying new problems independently, an expectation of being provided solutions, and an unwillingness to challenge the authority figure. Their observations are corroborated by research in younger children, such as that of Flores (2006, p.127). One potential solution, as observed by Campbell et al. (2020, p.764), is an increased focus on student collaboration within mathematics, to reduce the reliance on an external authority figure. In such instances, it would be interesting to investigate whether certain students are dominating discussions, and (if so) whether they have assumed the role of authority figures for the other students. Unfortunately, such an inquiry (as well as any subsequently suggested) was beyond the remit of this study.

Additionally, Harel and Sowder (1998, p.245) posited the existence of a *ritualistic* proof scheme, wherein students “believe that ritual and form constitute mathematical justification”. Furthermore, they suggest this may be a consequence of an “asymmetrical emphasis on teaching proof in algebra versus geometry” (ibid. p.234) in which students mostly encounter ideas of proof within geometry. This is corroborated by Bergwall (2021, p.746), who noted that Swedish steering documents mainly emphasise proof within geometry, trigonometry and number theory, and by Wells (2009), who criticised mathematics textbooks in England for avoiding the use of the word proof altogether in activities outside of geometry. Anecdotally, I have observed evidence of ritualistic proof schemes even within PGCE faculty sessions, such as the class collectively rejecting another student’s argument on the grounds that it was not sufficiently formal.

Finally, Harel and Sowder (1998, p.252) argued the existence of an *inductive* proof scheme, in which the validity of conjectures was ascertained by “quantitatively evaluating…in one or more cases”. The authors argue that this misconception is “natural” as “people’s evaluation of hypotheses in everyday life is probabilistic in nature”. Furthermore, they argue that the use of inductive reasoning is often “essential” for mathematical activities, and argue that mathematics education in school is “dominantly inductive”. This is supported by the National Curriculum (DfE, 2013, p.5), which demands the ability to “make and test conjectures” from all students (whilst deductive proof is non-statutory). Campbell et al. (2020, pp.758-759) observed similar expectations in the United States, whilst Miyazaki (2000) investigated strategies of moving between inductive and algebraic arguments.

Proof Conceptions in Algebra

Healy and Hoyles (2000) were also interested in investigating the difficulties students encountered with regards to proof. In particular, they criticised preceding research for theoretical (rather than empirical) approaches which did not pay sufficient attention to “documenting students’ views of mathematics” (ibid., p.297). To address this deficit, the authors conducted a large empirical survey of Year 10 students which investigated their perceptions of proof. Additionally, the authors conducted follow-up interviews with a small sample of respondents.

In total, the authors surveyed 2,459 students (1305 girls, 1154 boys) from 94 classes in 90 schools (Healy & Hoyles, 2000, p.405). This significantly exceeds the recommended minimum sample size of thirty cases recommended by Cohen et al. (2017, p.203), but more nuance is required. In particular, the authors make use of chi-squared statistics in their analysis (Healy & Hoyles, 2000, p.407), for which a minimum of five or more cases is recommended per point of comparison (Cohen et al., 2017, p.204). As the average number of students per class surveyed exceeds 25, this choice in test appears reasonable provided the authors did not stratify their dataset too much further.

Despite the size of the sample, there may yet be questions about its generalisability. This is because (similarly to the research of Harel and Sowder) it is focused on a subset of the student population of "higher than average ability" (Healy & Hoyles, 2000, p.405). For the authors, this was a conscious choice intended to guarantee the respondents had received "sufficient exposure" to the algebra involved in the survey. As in the work of Harel and Sowder, it further stands to reason that misconceptions of proof that appear in groups of students characterised by high mathematical attainment are likely to appear more generally.

An additional logistical concern is raised by the format of the survey – it is intended to be 70 minutes long, which exceeds the duration of many school lessons. Consequently, there is a risk of schools allotting different amounts of time to the survey, thereby weakening the strength of comparisons between different schools. This may not be an issue, as "a person employed by the research team oversaw the administration of the survey", but the extent of their involvement is unclear. In addition, the employment of an overseer for each of the 90 schools involved would likely incur significant additional costs.

The use of follow-up interviews to gather qualitative data enables the authors to investigate potential explanations for their quantitative data, and reflects an *explanatory sequential design* (Cohen et al., 2017, p.39). The number of interviews conducted is necessarily smaller – only 10 of the 90 schools – although it is unclear precisely how many students were interviewed in total. Additionally, these schools were sampled using a "multilevel analysis to identify schools that performed particularly well" (Healy & Hoyles, 2000, p.406). Consequently, their qualitative data may not be representative of their entire sample, and the addition of other groups – such as those which performed particularly poorly – might provide additional useful information.

As part of the quantitative survey, students were presented with a set of justifications of a given conjecture. They were then asked to select the justification which would be closest to their own, as well as the justification which they believed would receive the highest marks from their teacher. Notably, the results suggested a dichotomy between the two – "the arguments that were the most popular for the students' own approaches turned out to be the least popular when it came to choosing for best mark, and vice versa" (Healy & Hoyles, 2000, p.407). In other words, the students simultaneously held multiple conceptions of proof.

For example, whilst narrative arguments written in everyday language "were chosen by large numbers of students as closest to their own approach" (ibid., p.415), formal algebraic arguments (even those which were not valid) were the most popular when choosing for the best mark (ibid., 2000, p.412). Consequently, this research provides further quantitative evidence for the ritualistic proof scheme identified by Harel and Sowder (1998, p.245) described above.

Similarly, one of the justifications included was empirical in nature, which enables comparisons to the inductive proof scheme suggested by Harel and Sowder. In particular, it was found that whilst 24% of students chose the empirical argument as closest to their own, only 3% believed it would receive the best mark (Healy & Hoyles, 2000, p.406). Based upon this evidence, the authors argue that "most students were aware that empirical arguments had limitations".

However, this does not explain why empirical arguments "dominated responses" when students were asked to justify an unknown conjecture. The authors attribute this phenomenon to a skill deficit, arguing that they are "the best arguments available to the students" – and provide some qualitative data to support this (ibid., p.412). Similarly to Harel and Sowder, they attribute this to inductive practices within school mathematics, and in particular to the use of empirical data in investigations tasks, where "students are most likely to encounter reasoning and proof" (Healy & Hoyles, 2000, p.409).

Alternatively, one could argue that the disparities observed with respect to empirical arguments points to an awareness, rather than an acceptance, of the differing norms and values underlying students' and teachers' proof processes, and interpret this data using the frameworks of Dawkins and Weber

(2017) and Solomon (2006). One advantage of this argument is that it is supported by the high explanatory ratings afforded to empirical arguments by students – the authors found that "fewer than one third of the students felt that empirical argument had no explanatory value at all" (Healy & Hoyles, 2000, p.412).

Research Questions

In light of the above literature review, this paper will consider the following three Research Questions (RQs):

RQ1 – What do students believe a mathematical proof is? What forms of reasoning are involved? Do they perceive these tasks as valuable?

RQ2 – How do students approach tasks involving formal mathematical proof? Is there evidence of a skills deficit? How do students justify a result to themselves and to others?

RQ3 – Which misconceptions of proof can be observed within students' classroom work? Can they be related to the proof schemes identified by Harel and Sowder (1998)?

The Case Study

In order to investigate the above research questions, I performed a case study of a class of 30 high achieving year 10 students within my placement school. This study consisted of a sequence of three lessons regarding the topic of mathematical proof, each of which was approximately 55 minutes in length, as well as a homework assigned at the end of the third lesson. Subsequently, I conducted a series of interviews with students based upon their responses to both the lessons and the homework. In keeping with the British Educational Research Association guidelines (BERA, 2018, p.9), all students were informed of the purpose of the interviews, as well as their right to withdraw, in order to obtain informed consent. Additionally, this research adhered to the ethical procedures and practices of the Faculty of Education at the University of Cambridge, who (in conjunction with the school) acted as gatekeepers for the study.

The decision to undertake a qualitative case study is both practically and ethically motivated. Due to logistical constraints, this research could only be undertaken with one class of students. Consequently, there are too many issues with sample size and generalisability for this research to attempt any meaningful quantitative analysis. Furthermore, only a limited amount of time could be allocated to these lessons, so as to minimise impact on statutory teaching in adherence with educational research guidelines on “minimising any risk or harm to participants” (BERA, 2018, p.8). This makes any kind of action research inappropriate, as it would not be feasible to implement and evaluate any changes to practice across such a limited timeframe.

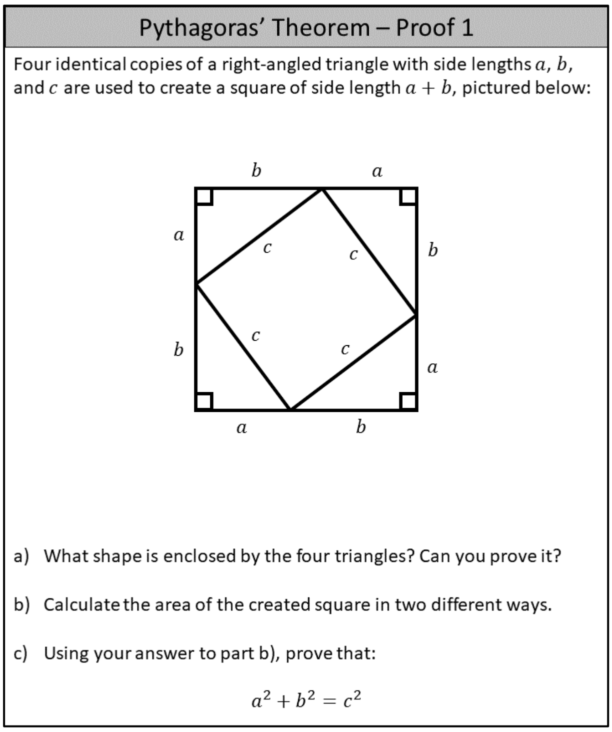
Instead, following the descriptions of Cohen et al. (2017, p.288), the design of this research is intended to “complement and augment” the data obtained by Healy and Hoyles (2000), to “provide alternative perspectives on topics” by interpreting them in the context of a wider literature review, and to “identify and refine questions” which may be worth further study.

Research Design

This case study can be broadly separated into two sections, which investigate proofs within the domains of geometry and algebra separately. This decision was informed by the differing levels of emphasis on proof in these two domains observed during the literature review by Harel and Sowder (1998), amongst others, and mirrors the decision of subsequent research – such as Healy and Hoyles (1998, 2000) to investigate proofs in both domains.

The first two lessons of this case study investigated mathematical proof within the context of geometry. During the first lesson, students were separated into groups of two to three members. Each group was then tasked with completing a particular proof of Pythagoras’ Theorem. In total, five different proofs were disseminated amongst the groups, with an additional sixth proof available as an extension task. An example of one of these tasks is illustrated in Figure 1 (next page).

Pythagoras’ Theorem was chosen as the object of proof for this lesson due to its familiarity, as well as for having a large number of proofs which were sufficiently accessible to students. The use of group work served the dual purpose of improving accessibility whilst reducing the need for teacher intervention, which facilitated more efficient data collection whilst also providing an opportunity to observe the ways in which students approached proof-related problems in the absence of an authority figure.



Pythagoras' Theorem – Proof 1

Four identical copies of a right-angled triangle with side lengths a, b, and c are used to create a square of side length $a+b$, pictured below:

a) What shape is enclosed by the four triangles? Can you prove it?

b) Calculate the area of the created square in two different ways.

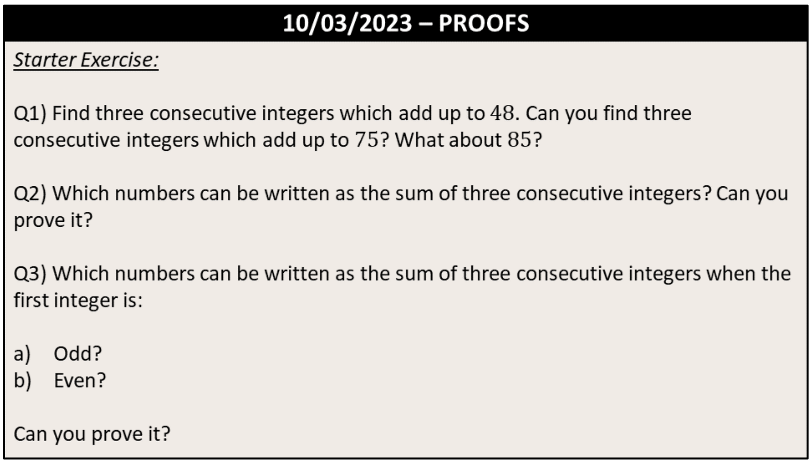
c) Using your answer to part b), prove that:

$$a^2 + b^2 = c^2$$

Figure 1: A scaffolded task in which students are required to complete a proof of Pythagoras' Theorem

During the second half of the first lesson, groups who had been working on the same task were matched up in order to discuss their proofs and to prepare a presentation for the second lesson, in which they would present their arguments to the rest of the class. At the end of each presentation, students were encouraged to ask questions in an effort to further uncover any difficulties or misconceptions they had encountered. This presented an opportunity for a detailed observation of the ways in which students justify an argument to themselves and others, whilst simultaneously connecting to a schoolwide target of improving student oracy.

The final lesson of the case study focused on algebraic proof, and may be broken down into several sections. In the first section, students were tasked with investigating which numbers could be written as the sum of three consecutive integers, leading to further discussions based on the parity of the first integer. This is illustrated in Figure 2 (next page).



10/03/2023 – PROOFS

Starter Exercise:

Q1) Find three consecutive integers which add up to 48. Can you find three consecutive integers which add up to 75? What about 85?

Q2) Which numbers can be written as the sum of three consecutive integers? Can you prove it?

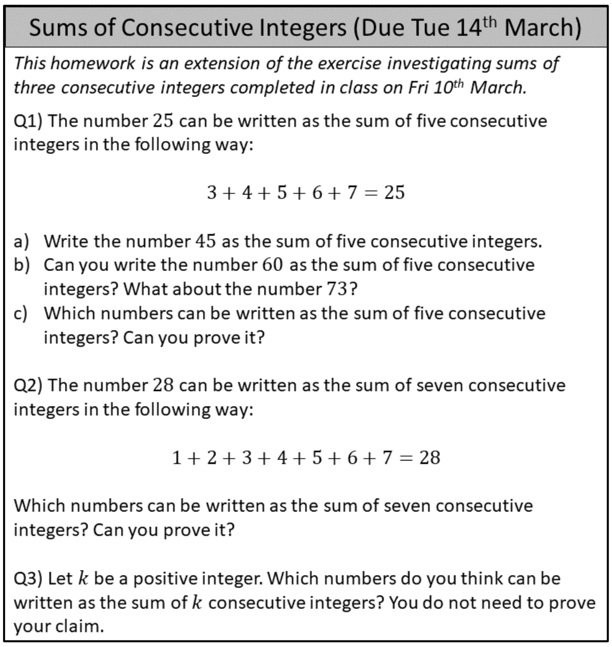
Q3) Which numbers can be written as the sum of three consecutive integers when the first integer is:

a) Odd?
b) Even?

Can you prove it?

Figure 2: A task in which students investigate the sums of three consecutive integers

This task was chosen for two main reasons. Firstly, it is accessible; students encounter algebraic representations of sums of consecutive integers as early as year 7 within my placement school. Secondly, it is generalisable. For example, one may distinguish between sums of consecutive integers based on the parity of the starting integer, thus providing a context in which to discuss algebraic representations of odd and even integers. Alternatively, one may investigate the sum of an arbitrary number k of consecutive integers. This latter generalisation forms the basis of an accompanying homework piece seen in Figure 3.



Sums of Consecutive Integers (Due Tue 14th March)

This homework is an extension of the exercise investigating sums of three consecutive integers completed in class on Fri 10th March.

Q1) The number 25 can be written as the sum of five consecutive integers in the following way:

$$3 + 4 + 5 + 6 + 7 = 25$$

a) Write the number 45 as the sum of five consecutive integers.
b) Can you write the number 60 as the sum of five consecutive integers? What about the number 73?
c) Which numbers can be written as the sum of five consecutive integers? Can you prove it?

Q2) The number 28 can be written as the sum of seven consecutive integers in the following way:

$$1 + 2 + 3 + 4 + 5 + 6 + 7 = 28$$

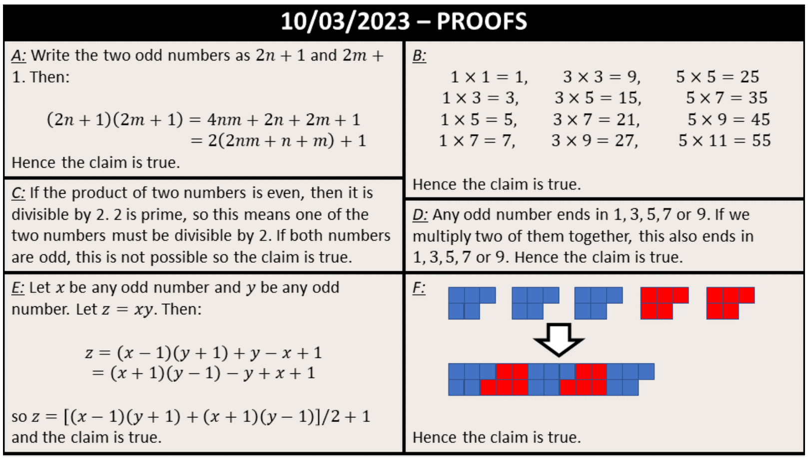
Which numbers can be written as the sum of seven consecutive integers? Can you prove it?

Q3) Let k be a positive integer. Which numbers do you think can be written as the sum of k consecutive integers? You do not need to prove your claim.

Figure 3: Homework in which students investigate the sum of k consecutive integers

Notably, students only investigate sums of an odd number of consecutive integers before conjecturing a general result. This choice is deliberate; the sums of k consecutive integers correspond to multiples of k precisely when k is odd. Consequently, this presented a novel opportunity to investigate the perils of empirical reasoning.

During the remainder of the lesson on algebraic proofs, students spent some time recording their beliefs on proof before investigating the statement "the product of two odd numbers is odd". After being given some time to attempt to prove the statement themselves, students were asked to evaluate the justifications presented below in Figure 4.



10/03/2023 – PROOFS

A: Write the two odd numbers as $2n+1$ and $2m+1$. Then:

$$\begin{aligned}(2n+1)(2m+1) &= 4nm+2n+2m+1\\ &= 2(2nm+n+m)+1\end{aligned}$$

Hence the claim is true.

B:

$$\begin{array}{ccc} 1\times 1 = 1, & 3\times 3 = 9, & 5\times 5 = 25\\ 1\times 3 = 3, & 3\times 5 = 15, & 5\times 7 = 35\\ 1\times 5 = 5, & 3\times 7 = 21, & 5\times 9 = 45\\ 1\times 7 = 7, & 3\times 9 = 27, & 5\times 11 = 55\end{array}$$

Hence the claim is true.

C: If the product of two numbers is even, then it is divisible by 2. 2 is prime, so this means one of the two numbers must be divisible by 2. If both numbers are odd, this is not possible so the claim is true.

D: Any odd number ends in 1, 3, 5, 7 or 9. If we multiply two of them together, this also ends in 1, 3, 5, 7 or 9. Hence the claim is true.

E: Let x be any odd number and y be any odd number. Let $z = xy$. Then:

$$\begin{aligned}z &= (x-1)(y+1)+y-x+1\\ &= (x+1)(y-1)-y+x+1\end{aligned}$$

so $z = [(x-1)(y+1)+(x+1)(y-1)]/2+1$ and the claim is true.

F:

Hence the claim is true.

Figure 4: A set of justifications for the statement "the product of two odd numbers is odd"

This task is heavily inspired by the work of Healy and Hoyles (2000) discussed in the literature review. Similarly, the students in this task were asked to evaluate the six justifications from a teacher's perspective as well as their own. These justifications include an algebraic proof (A), an empirical evaluation (B), a narrative proof (D) and a visual argument using a generic example (F). Justification E is an incorrect algebraic response intended to assess whether students were applying ritualistic reasoning, but differs from those used by Healy and Hoyles in that the algebraic manipulations are not incorrect – they just do not prove the conjecture in question. Justification C is a second narrative proof which enables further investigation into the assertion that students are not convinced by proof by contradiction (Harel & Sowder, 1998, p.253).

Additional follow-on tasks involving squares of odd numbers and square roots of even square numbers were planned for the lesson and had been intended to investigate whether students could recognise a specific instance of a more general argument as well as to explore the contrapositive. Due to time constraints, these were omitted.

Data Collection

To conclude this section, I will discuss the types of data which were collected during this case study, and highlight some of the advantages and disadvantages of each. Additionally, I will briefly indicate how each type of data was used in relation to the research questions.

The first main source of data collected for this case study consists of observation notes for each lesson, written by both the regular class teacher and myself. Having a second observer was particularly beneficial here, as it helped to mitigate any potential attention deficit (Cohen et al., 2017, p.560) necessitated by simultaneously teaching the class. Moreover, the second observer's familiarity with the class – as well as the frequency of departmental observations within the school - meant that their presence as a second observer was less likely to cause reactivity amongst the students (Cohen et al., 2017, p.560), thereby reducing the risk of bias. These notes were subsequently transcribed and coded according to the conceptions of proof identified by Harel and Sowder (1998). In order to avoid problems of inference (Cohen et al., 2017, p.561), this coding focuses solely on identified misconceptions, rather than their underlying causes.

Another important source of data for this case study consists of scans of the classwork completed by the students during the lessons, as well as the homework. One advantage of this data is that it was very simple to collect outside of lessons, and thus did not create additional logistical difficulties. However, it shares similar problems of inference to the aforementioned observational data. Additionally, I encountered some issues of non-response from students in both classwork and homework, as well as instances of destruction of data when students worked on mini-whiteboards.

For logistical reasons, students' classwork was not transcribed and coded, although student evaluations of the six proofs from the third lesson of the case study were tabulated. The purpose of this tabulation was not to draw quantitative conclusions regarding misconceptions of proof, as the sample size involved is far too small. Instead, it was used (in conjunction with the observational data and classwork) to identify which students to interview.

In total, two pairs of students (henceforth referred to as Students A, B, C and D) were interviewed for approximately half an hour each. These interviews were conducted away from the main lesson, so the use of paired interviews ensured that appropriate ethical guidelines were being followed. Furthermore, students with contrasting answers were selected to encourage them "to challenge each other and participate in a way that may not happen in a one-to-one, adult-child interview" (Cohen et al., 2017, p.529). For ethical reasons, these interviews were not recorded. Instead, they were transcribed and subsequently coded in the same way as the observation notes.

The following two-way table (Table 1) illustrates which of the three data types were used to answer each of the three Research Questions.

	Observation Notes	**Students Work**	**Interviews**
Research Question 1	Yes	No	Yes
Research Question 2	Yes	Yes	No
Research Question 3	Yes	Yes	Yes

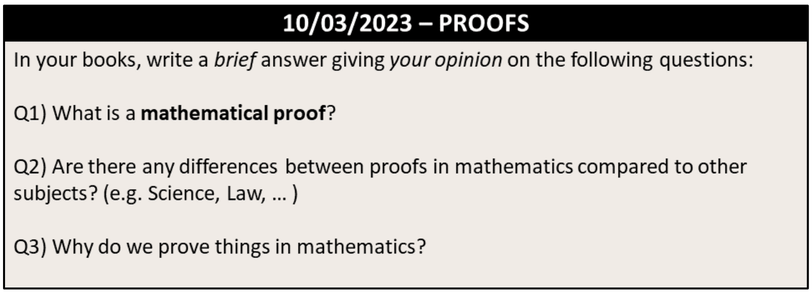
Table 1: Which research questions were investigated using which data types

Results

The second half of this paper is dedicated to an analysis of the results obtained during the case study, which is structured so as to address each of the research questions in turn. To begin with, I will discuss any results pertaining to students' perceptions of mathematical proof itself. After this, I will discuss how the students interacted with the various kinds of mathematical proof encountered in the case study, and what implications this might have for the teaching of proof. Finally, I will discuss any data related to each of the specific misconceptions of proof identified by Harel and Sowder (1998) in more detail.

Perceptions of Proof

During the third lesson of the case study, students were asked the three questions seen below in Figure 5 (next page).



10/03/2023 – PROOFS

In your books, write a *brief* answer giving *your opinion* on the following questions:

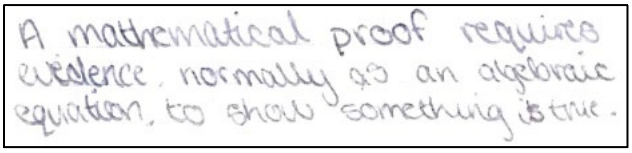
Q1) What is a **mathematical proof**?

Q2) Are there any differences between proofs in mathematics compared to other subjects? (e.g. Science, Law, …)

Q3) Why do we prove things in mathematics?

Figure 5: A task in which students were asked to record their views on mathematical proof

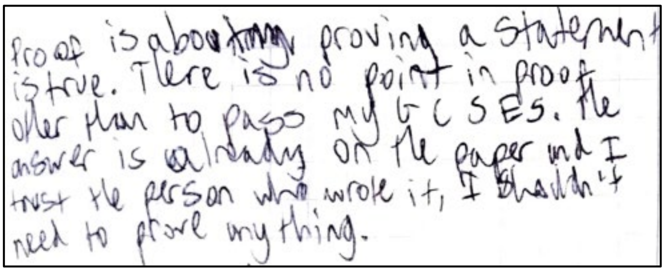
These questions were notable casualties of non-response, with less than half of the students providing an answer for any of the questions asked. One possible explanation of this would be a lack of confidence on the part of students regarding the nature of proof, which would be consistent with the observations of Dawkins and Weber (2017, p.134) and Healy and Hoyles (2000, p.418), but this inference requires more data to justify. The majority of the students who did respond focused on proof as a means of justification, see Figure 6 for one example, which is again consistent with Healy and Hoyles (2000, p.418).



A mathematical proof requires
evidence, normally as an algebraic
equation, to show something is true.

Figure 6: A student's response centred around the use of proof for justification

A small number of students described alternative uses of proof. For example, Student A afforded proof an explanatory function, arguing that "You might want to prove something to see how you get there – to extend your understanding". Student B went even further, arguing that proof could be used "to form a larger, more complicated theory". On the other hand, some students argued that proof had no function at all, as illustrated by Figure 7.



Proof is about proving a statement
is true. There is no point in proof
other than to paso my GCSEs. The
answer is already on the paper and I
trust the person who wrote it, I shouldn't
need to prove anything.

Figure 7: A student's response which debated the merits of proof

Students' perspectives on the value of proof were similarly divided. Those students who questioned the value of proof were generally the most negative, with one student arguing "I don't think there's a point. It's very boring". Others viewed it as a problem-solving exercise, with Student D arguing that "it was quite complicated" but that "once we got it, it was enjoyable". The most positive responses came from Students A and B, who argued that they found it enjoyable to "see why things work the way they work" and likened it to their curiosities for mechanics and magic tricks, respectively.

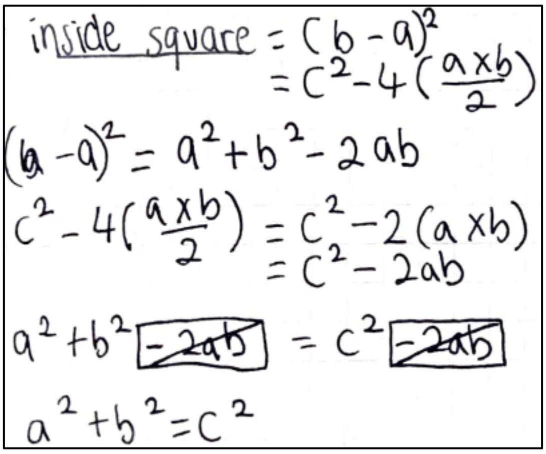
Notably, these results suggest a potential link between students' understanding of the various functions of proof and their enjoyment of proof itself, which merits further investigation. If such a correlation exists (which I do not claim here, as the data is far too limited) then this may be useful in informing future pedagogical practice by e.g. placing greater emphasis on the explanatory functions of mathematical proof.

Students' Approaches to Mathematical Proof

Setting aside debates on the function of proof, some authors such as Dawkins and Weber (2017) have argued that students "at all levels" have difficulty proving things altogether. This case study afforded the opportunity to investigate such claims through a variety of proof-related activities.

The first two lessons of the case study, in which students investigated Pythagoras' Theorem, were encouraging. By the end of the first lesson, all students had successfully completed a proof of the theorem. Moreover, the majority of errors that were made along the way could be categorised as algebraic, such as the incorrect expansion $(b-a)^2 = b^2 - a^2$, rather than conceptual. Conversely, many students demonstrated a very strong proficiency in handling the underlying algebra, with two students successfully completing a sixth extension proof which demanded an even greater level of algebraic mastery, as illustrated in Figure 8 (next page).

Notably, several students made the decision to first investigate the proof in the special case of the 3:4:5 triangle. This could represent either an algebraic misconception – that of premature evaluation, identified by Hart (1982, p.105) – or an instance of inductive reasoning. However, student responses – that it was "easier to do it for numbers" and that they "could repeat it for any number" – suggest that they instead used the 3:4:5 triangle as a generic counterexample.



inside square $= (b-a)^2$

$= c^2 - 4\left(\frac{a \times b}{2}\right)$

$(b-a)^2 = a^2 + b^2 - 2ab$

$c^2 - 4\left(\frac{a \times b}{2}\right) = c^2 - 2(a \times b)$

$= c^2 - 2ab$

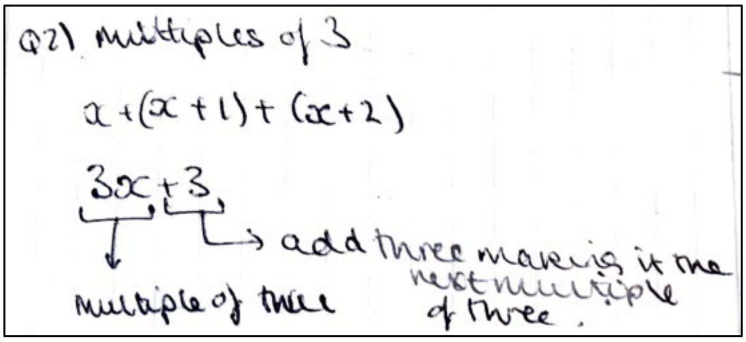
$a^2 + b^2 \boxed{\cancel{-2ab}} = c^2 \boxed{\cancel{-2ab}}$

$a^2 + b^2 = c^2$

Figure 8: A student's response to Proof 5 of Pythagoras' Theorem

Whilst students were adept at completing proofs themselves, they subsequently struggled to justify their arguments to others. Usually, this was due to omission of key details. For example, the students presenting Proof 2 did not differentiate between which similar triangles they were given and which had been constructed, which resulted in confusion amongst the audience about what was being assumed and what was being proven. Indeed, when asked about her favourite proof of Pythagoras' theorem, Student B responded that she preferred her own, as "the students weren't teachers and that made the explanation harder".

In the final lesson, students encountered an additional complexity: they were now required to formulate their arguments independently, whereas their proofs of Pythagoras' Theorem had been carefully scaffolded. For the starter task shown in Figure 2, this did not appear to be an issue, as students had likely encountered algebraic representations of consecutive integers beforehand (see example in Figure 9).



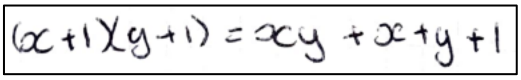
Q2) multiples of 3

$x + (x+1) + (x+2)$

$\underbrace{3x}_{\text{multiple of three}}\underbrace{+3}_{\text{add three making it the next multiple of three.}}$

Figure 9: A student's argument that sums of three consecutive integers are always multiples of 3

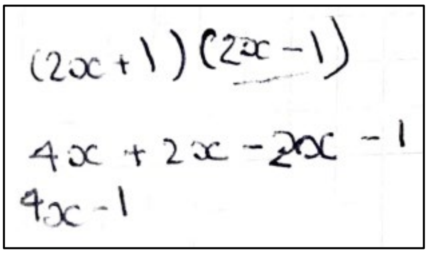
Conversely, this lack of scaffolding caused major issues when students were invited to prove that the product of two odd numbers is odd, as they were unsure how to represent the two numbers algebraically. This led to a large variety of misconceptions. For example, many students were unsure how to encode the information that the numbers were odd, as illustrated in Figure 10.



$(x+1)(y+1) = xy + x + y + 1$

Figure 10: A student's response which does not successfully encode the parity of the two numbers

In other cases, students could represent odd numbers algebraically, but attempted to do so using a single variable. This resulted in a number of proofs of very specific subcases, such as when the two numbers are consecutive (as seen in Figure 11).



$(2x+1)(2x-1)$

$4x + 2x - 2x - 1$

$4x - 1$

Figure 11: A student's response which mistakenly assumes the numbers are consecutive

This can be likened to the results obtained by Healy and Hoyles (2000, p.414), whose analysis of students' attempts to prove an unfamiliar conjecture concludes that "our students have difficulty presenting arguments algebraically". When compared to their success at proving Pythagoras' theorem in the first two lessons, this highlights the potential pedagogical importance of appropriate scaffolding so that students can focus on the reasoning aspects of proof-related tasks.

Misconceptions and Proof Schemes

In this section, I will analyse results from the case study which pertain to the system of proof schemes identified by Harel and Sowder (1998). In particular, I will focus on the three proof schemes (*authoritarian, ritualistic*, and *inductive*) which were described during the literature review. Additionally, I will discuss a special type of empirical argument - described by Harel and Sowder (1998, p.255) as a *perceptual proof scheme* – which became prominent over the course of the case study.

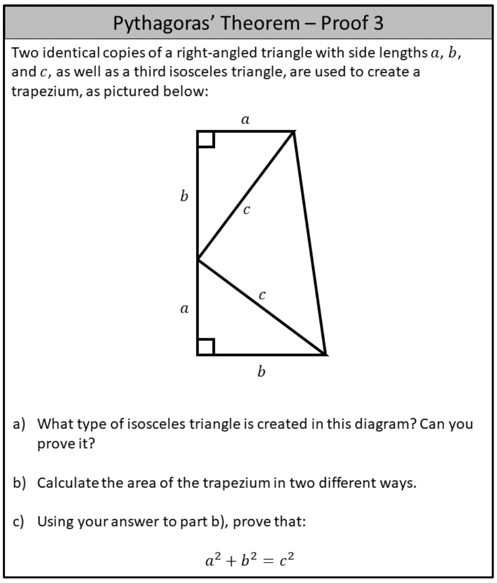
Authoritarian Proof Schemes

If a student's main source of conviction for a mathematical result is an external source, such as a textbook or a teacher, then they are said to possess an authoritarian proof scheme. I encountered multiple potential instances of this type of reasoning during the case study. Indeed, one particularly prominent example has already been illustrated in Figure 7, wherein a student argued that "the answer is already on the paper and I trust the person who wrote it, I shouldn't need to prove anything".

During the first two lessons, the use of group work provided an opportunity to investigate an assertion by Campbell et al. (2020, p.764) that an "increased focus on student collaboration" could be used to reduce the impact of authoritarian proof schemes. In particular, I wanted to discover whether the more authoritative students became external sources of conviction for the members of their group.

The results of this investigation were inconclusive. On one hand, certain students quickly emerged as "leaders" within their own groups. This was especially noticeable during presentations, as some groups deferred the task of presenting to a single student whilst another group struggled to reconstruct the proof in the absence of a key member. On the other hand, neither of these episodes provide sufficient evidence to infer that those students were acting as an external source of conviction for the other students in their group.

A more compelling argument for students' use of authoritarian proof schemes can be made based on their presentation of Proof 3 of Pythagoras' Theorem, presented in Figure 12.



Pythagoras' Theorem – Proof 3

Two identical copies of a right-angled triangle with side lengths a, b, and c, as well as a third isosceles triangle, are used to create a trapezium, as pictured below:

a) What type of isosceles triangle is created in this diagram? Can you prove it?

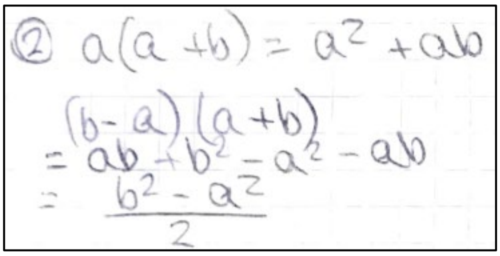
b) Calculate the area of the trapezium in two different ways.

c) Using your answer to part b), prove that:

$$a^2 + b^2 = c^2$$

Figure 12: Proof 3 of Pythagoras' Theorem, based on the proof by President James Garfield

As part of this proof, students were required to compute the area of a trapezium, which they chose to do by dissecting the trapezium into a rectangle and a triangle, as illustrated in Figure 13.



② $a(a+b) = a^2 + ab$

$(b-a)(a+b)$

$= ab + b^2 - a^2 - ab$

$= \dfrac{b^2 - a^2}{2}$

Figure 13: A student's computation of the area of the trapezium using the dissection described above

Notably, this dissection depends on the assumption that the side lengths of the triangle satisfy the inequality $a < b$. When another student questioned this assumption, the presenting group claimed they were told it was true. This answer was deemed satisfactory by the rest of the class, who did not ask any further questions. However, the presenting group had been told nothing of the sort! This has the concerning implication that students may not even need to verify external sources of information before using them as a basis for their convictions.

Another striking encounter indicative of an authoritarian proof scheme took place during my interview with Students A and B. As part of this interview, I was particularly interested in investigating their assessments of the justifications shown in Figure 4 (earlier) of the algebraic statement that "the product of two odd numbers is odd". See Figure 14 for Student A's response.

	me	teacher	friend
A)	✓	✓	✓
B)	✓	✗	✓
C)	✗	✓	✗
D)	✓	✓	✓
E)	✓	✓	✗
F)	✓	✓	✓

Figure 14: Student A's assessment of the arguments shown in Figure 4

Recall that Proof E was a false algebraic proof, intended to identify instances of ritualistic proof schemes. When questioned why she had accepted the argument, Student A said that she could not remember. When pressed for details, she argued that it "made sense that the claim would be true", before finally claiming that she had been convinced by the argument "because you gave it to us as a teacher". This appeal to authority was especially surprising in light of Student A's evaluation of

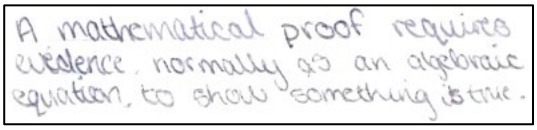
Proof B in Figure 14 (above), which demonstrated an awareness that not all of the arguments would be accepted by the teacher.

The two examples of authoritarian arguments discussed so far contain a key similarity – they both occurred when a student was asked to justify a statement under pressure. Consequently, one possible explanation for their appeal to an authoritative proof scheme is that – in the moment, owing to the pressure – it was the only proof scheme available. This interpretation is supported by Harel and Sowder (1998, p.254), who suggested more generally that students justify their conjectures using authoritative (and empirical) proof schemes because "these are the only proof schemes they possess". This stresses the importance of teaching students how to prove mathematical statements, as well as the benefits of ensuring that students are given sufficient time when engaged in proof-related activities.

Ritualistic Proof Schemes

Another proof scheme encountered during this case study was the ritualistic proof scheme. This proof scheme was introduced by Harel and Sowder (1998) and may be categorised by an emphasis on the form or appearance of an argument over its content.

In this case study, the ritualistic proof scheme appeared to manifest itself as an extreme preoccupation with algebra and symbolism. Indeed, some students embedded algebra into their very definition of proof, as exemplified in Figure 15.



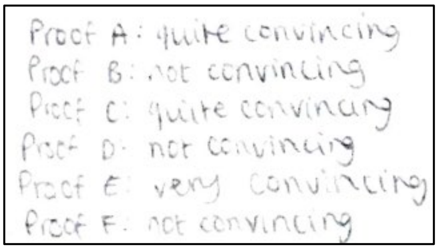
A mathematical proof requires evidence, normally as an algebraic equation, to show something is true.

Figure 15: A student's definition of proof which explicitly references algebra

This observation is supported by data from both interviews. During the first, a discussion led by Student A about the importance of generality within proof prompted a response of "like algebra" from Student B, which was subsequently endorsed by Student A. The second interview was even more explicit; Student D defined mathematical proof as "having to prove something with algebra, facts".

There were two apparent consequences of the above coupling between algebra and proof. Firstly, students had misgivings about arguments which were not presented algebraically. This is corroborated by Harel and Sowder (1998, p.246), who observed that students often had doubts about “whether a certain justification is considered a proof” when the argument in question “was not communicated via mathematical notations”. As a result, students exhibited a clear preference for algebraic arguments. Indeed, when invited to prove the statement “the product of two odd numbers is odd”, every single recorded response was an attempt at an algebraic proof. This represents a significant departure from a similar activity conducted by Healy and Hoyles (2000, p.408), in which only 11% of students attempted to prove a conjecture using algebra. Their rate of non-response (3%) was also much lower, which may not be surprising in light of earlier discussions of the difficulties students face when presenting arguments algebraically.

At a more extreme level, this preference for algebraic arguments resulted in the rejection of arguments which did not “appear mathematical” but were otherwise correct. This could be observed in some of the students’ evaluations of the proofs shown in Figure 4.



Proof A: quite convincing
Proof B: not convincing
Proof C: quite convincing
Proof D: not convincing
Proof E: very convincing
Proof F: not convincing

Figure 16: Student D’s evaluations of the arguments presented in Figure 4

Student D’s responses to this task (shown above in Figure 16) were particularly interesting, as she accepted only one argument – Proof C – which was not algebraic. When questioned about this proof in particular, Student D argued that it was “quite strategic” and “uses stuff we use every day so it must be true”. More precisely, she appeared to be referring to the use of prime numbers and divisibility within the proof. This is arguably consistent with a ritualistic proof scheme, as it is possible that Student D’s response was based solely on the appearance of mathematical vocabulary, rather than its application. When asked why she had rejected Proof D, Student D’s response was more straightforward: she did not like it “because it stated the obvious…it’s not really showing any algebra to prove the point”. These misgivings about the lack of algebra were shared by other interviewees; Student B felt that the proof was “too simple” whilst even Student A – who did not reject the proof, as seen in Figure 14 – suggested that it was “more suited to younger years who don’t understand the algebra”.

The second apparent consequence of the ritualistic proof scheme is the acceptance or verification of an incorrect proof due to its ritualistic aspects. This is not a new phenomenon; it has been observed in several earlier publications, such as Martin and Harel (1989), Harel and Sowder (1998) and Healy and Hoyles (2000). As explanation, the authors of the second paper suggest that it may be attributed to an “over-emphasis in schools of proof writing prior to and even in place of proof understanding, production and appreciation” (Harel & Sowder, 1998, p.246). This appears to be a reference to the “two-column proof” taught in geometry classes in the USA, which was discussed earlier in the same paper (Harel & Sowder, 1998, p.236) and which has been investigated by other authors for, amongst other things, “perpetrating a formalistic image of proving” (Weiss et al., 2009, p.275). As this argument appears to be based on an entirely different curriculum, it should not be applied to this case study. Instead, I would like to use the data collected here to propose a different – and more general – explanation.

The data in question is derived from a disagreement between Student C and Student D regarding Proof E, the false algebraic justification given in Figure 4 of the statement “the product of two odd numbers is odd”. Initially, Student D argued that Proof E was “very convincing” because “it had lots of algebra”. This did not satisfy Student C, because the argument itself did not make sense to her. When this objection was raised, Student D admitted that it did not make sense to her either. When asked why she had accepted the argument anyway, her response was very revealing: “I just thought I was too dumb that I wouldn’t understand it”.

Based upon the above episode, I would like to propose that ritualistic proof schemes may have an authoritative explanation: when students cannot understand a proof that looks compelling, they assume that the proof is correct and that they are wrong. This would be consistent with the authoritative response given to the task by Student A, and suggests a potential link between ritualistic proof schemes and maths anxiety which may merit further research.

Empirical Proof Schemes

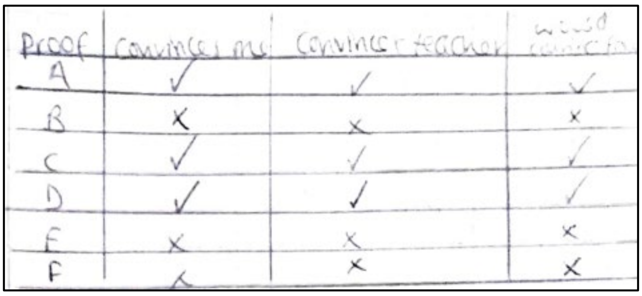
Finally, I will discuss data from the case study relating to *empirical* proof schemes, in which conjectures are validated “by appeals to physical facts or sensory experiences” (Harel & Sowder, 1998, p.252). Following their classification, this section will be divided into two sections, focusing on *inductive* and *perceptual* proof schemes, respectively.

An *inductive* proof scheme arises when a student validates a mathematical statement by evaluating it quantitatively. Harel and Sowder (1998, p.252) argue that this type of reasoning is “dominant” among students. As evidence, they cite an earlier paper by Martin and Harel (1989) which found that “more than 80% of 101 prospective teachers considered inductive arguments to be mathematical proofs”, as well as a dissertation by Goetting (1995) which found that “almost 40% of her advanced undergraduates” employed an inductive proof scheme. Healy and Hoyles (2000, p.396) expand upon this further, suggesting that while empirical arguments “predominate in students’ own proof constructions”, this was a consequence of a skill deficit, and “most students were aware of its limitations”.

The findings from this case study appear to support the latter viewpoint, with most students’ rejecting the inductive argument – Proof B – given in Figure 4. Student A’s response (see Figure 14) stood out here, as she did not reject the argument despite recognising that a teacher would, which corroborates the claim by Healy and Hoyles (2000, p.396) that students may hold multiple conceptions of proof. Her interview responses, however, were much more one-sided: she argued that choosing lots of examples “doesn’t count, you’ve got to have a formula”, and supported this viewpoint by reasoning that “If I told you every multiple of 3 was a multiple of 6 and showed you all the multiples of 6 that wouldn’t show you that it worked”. This apparent contradiction may exemplify a particular difficulty with interviewing children observed by Cohen et al. (2017, p.530); in this case, it appears that Student A gave the response “they think the interviewer wants to hear”.

Although students tended to reject inductive arguments overall, this was contingent upon the recognition that an argument was inductive to begin with. Based upon the results of the case study, this was not always clear, particularly when the arguments involved were visual proofs or generic examples.

As seen in Figure 17 (next page), Student C rejected Proof F, which used a visual proof of a generic example to justify the statement “the product of two odd numbers is odd”. When asked why, she argued that “It didn’t have any explanation. It’s just showing 5 lots of 5”. Student D concurred, stating that “it’s only one pair of odd numbers so it’s not really a proof”, although when asked to generalise it she suggested that “we can fit blocks together and if there’s an odd it’ll always have one left over”.



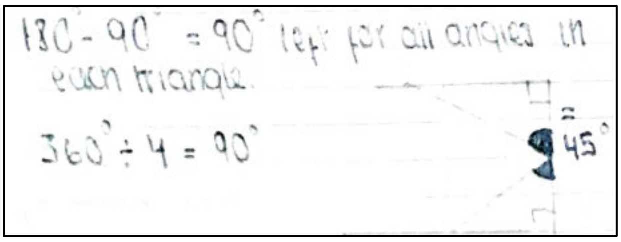
Proof	Convinces me	Convinces teacher	would convince fr
A	✓	✓	✓
B	✗	✗	✗
C	✓	✓	✓
D	✓	✓	✓
E	✗	✗	✗
F	✗	✗	✗

Figure 17: Student C's evaluations of the arguments presented in Figure 4

Conversely, Student B immediately recognised the generalisable nature of the proof, saying that "it's an example, but it's an example that applies to multiple situations. Every odd number has a sticky out bit, they fit together, and you'll have one left over". On the other hand, when asked for their favourite proof of Pythagoras' theorem, Student's A and B referenced a visual demonstration of the proof involving water without recognising its inductive nature.

Another interesting aspect of inductive reasoning was revealed by Student C's rejection of Proof B, as it appeared that she did not reject the inductive argument based solely on the fact it consisted of examples. Instead, she rejected it on the basis that "It gave you the first three odd numbers, it's not random numbers" before adding that "If the numbers were more random like 501 times something I'd believe it more". Similarly, when appraising Proof D, Student C appeared to reference "random" numbers, stating that "if you took 11 times something then it's an example of it being used". This may be linked to an observation by Harel and Sowder (1998, p.254) that students often verify statements using a *single* example, rather than multiple. In order to explain this, they argue that "the students believed that since the single example was chosen randomly and it conformed to the general statement, the general statement must be true". This misconception may arise from the use of probability to describe a general statement – "it works for any random number", for example – and has pedagogical implications regarding the use of language and its unexpected long-term impacts.

Finally, I would like to briefly discuss *perceptual* proof schemes, which are a subclass of empirical proof schemes in which students evaluate an argument based upon a specific visual representation without considering any possible variations. These proof schemes appeared prominently during the lessons involving Pythagoras' theorem – one example, in which students assumed the ordering $a < b$ of side lengths in order to enable a specific dissection, has already been discussed. Another example, pictured in Figure 18 (next page), shows students explicitly evaluating a pair of angles based upon the specific diagram they had drawn.



180 - 90° = 90° left for all angles in each triangle.

360° ÷ 4 = 90°

Figure 18: A student evaluates the angles in a general right-angled triangle

In the most extreme case, a group of students used a ruler to measure the side lengths of the triangles in a diagram they were given. Notably, this has significant parallels with the algebraic misconception of premature evaluation investigated by Hart (1982, p.105), which may merit further investigation.

Conclusion

Whilst proof holds a special place within mathematics, its place within the wider spectrum of mathematics education is altogether less certain. Indeed, mathematical proof remains a largely non-statutory part of the national curriculum, despite objections from academic circles which cite a lack of previous experience with proof as a major roadblock when encountering degree level mathematics. These claims are supported by a variety of educational research, and the difficulties encountered by undergraduates and high-school students regarding mathematical proof are well documented. Even if one were to momentarily disregard any questions regarding the difficulty of proof, questions of its importance remain, and it is far from a given that proof is as central to mathematics to most people as it is to mathematicians.

Earlier, it was stated that the goal of this case study was to complement and augment the data obtained by wider pieces of educational research, to provide alternative perspectives on topics by interpreting them in the context of a wider literature review, and to identify and refine questions which may be worth further study. In this regard, I feel this case study has been successful. Indeed, three of the key observations and questions raised are summarised below:

1) Students who held more sophisticated views of the purposes of proof appeared to enjoy it more. Is this correlation supported by quantitative research? If so, does it imply a causation, which could be used to improve student motivation towards proof-related exercises?

2) Are ritualistic proof schemes a result of an underlying authoritative assumption that the inability to comprehend a proof must be a deficit on the part of a student? If so, are they related to broader phenomena such as maths anxiety?

3) To what extent do students conflate concepts of generality and randomness? In particular, does this inform the development of inductive proof schemes?

However, there is a cruel irony to this case study – whilst mathematical proof is often viewed as absolute, these observations are the opposite, based upon a single author's interpretation of a minuscule sample size and thereby prone to all sorts of bias and sampling error. Nevertheless, they are largely consistent with the existing literature surrounding mathematical proof and may prove fertile ground for future – and more substantive – research.

On a more personal level, I am keen to further explore proof-related tasks in the classroom. This has certainly been encouraged by this case study, which yielded the key pedagogical insight that students were much more successful at proof-related tasks when they were scaffolded as problem-solving exercises. This allowed the students to focus on individual items of reasoning without introducing external difficulties such as algebraic formulation. Indeed, the key takeaway from this case study may be the observation that problem-solving exercises which constitute individual steps of a scaffolded mathematical proof appear to provide an accessible entry point from which students can use statutory knowledge to meaningfully engage with the topic.

References

British Educational Research Association (BERA) (2018). *Ethical guidelines for educational research.* BERA.

Bergwall, A. (2021). Proof-related reasoning in upper secondary school: characteristics of Swedish and Finnish textbooks. *International Journal of Mathematical Education in Science & Technology*, *52*(5), 731–751.

Campbell, T. G., Boyle, J. D., & King, S. (2020). Proof and argumentation in K-12 mathematics: a review of conceptions, content, and support. *International Journal of Mathematical Education in Science & Technology*, *51*(5), 754–774.

Cohen, L., Manion, L., & Morrison, K. (2017). *Research Methods in Education* (8th. ed.). London: Routledge.

Dawkins, P., & Weber, K. (2017). Values and norms of proof for mathematicians and students. *Educational Studies in Mathematics*, *95*(2), 123–142.

Department for Education (DfE) (2013). *Mathematics: programmes of study: key stage 4*. DfE. Retrieved from https://www.gov.uk/government/publications/national-curriculum-in-england-mathematics-programmes-of-study

Flores, A. (2006). How do students know what they learn in middle school mathematics is true? *School Science and Mathematics*, *106*(3), 124–132.

Goetting, M. (1995). The college students' understanding of mathematical proof (Doctoral dissertation, University of Maryland). *Dissertations Abstracts International*, 56, 3016A.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Issues in mathematics education: Vol. 7. Research in collegiate mathematics education III* (pp. 234–283). American Mathematical Society.

Hart, K. (1982). *Children's Understanding of Mathematics 11-16.* John Murray.

Healy, L., & Hoyles, C. (1998). *Justifying and proving in school mathematics* (Technical Report). Institute of Education, University of London.

Healy, L., & Hoyles, C. (2000). A Study of Proof Conceptions in Algebra. *Journal for Research in Mathematics Education*, *31*(4), 396–428.

London Mathematical Society. (1995). *Tackling the Mathematics Problem.* LMS.

Martin, W. G., & Harel, G. (1989). Proof Frames of Preservice Elementary Teachers. *Journal for Research in Mathematics Education*, *20*(1), 41–51.

Miyazaki, M. (2000). Levels of Proof in Lower Secondary School Mathematics. *Educational Studies in Mathematics*, *41*(1), 47-68

Solomon, Y. (2006). Deficit or difference? The role of students' epistemologies of mathematics in their interactions with proof. *Educational Studies in Mathematics*, *61*(3), 373–393.

Weiss, M., Herbst, P., & Chen, C. (2009). Teachers’ Perspectives on “Authentic Mathematics” and the Two-Column Proof Form. *Educational Studies in Mathematics*, *70*(3), 275–293.

Wells, D. (2009). Where’s the Proof?. *Mathematics in School*, *38*(5), 27-29.