Doing a research project: the messiness of qualitative research

an example from the work with Interactive Theorem Provers

Paola Iannone

The University of Edinburgh

paola.iannone@ed.ac.uk

THE UNIVERSITY *of* EDINBURGH

LEAN THEOREM PROVER

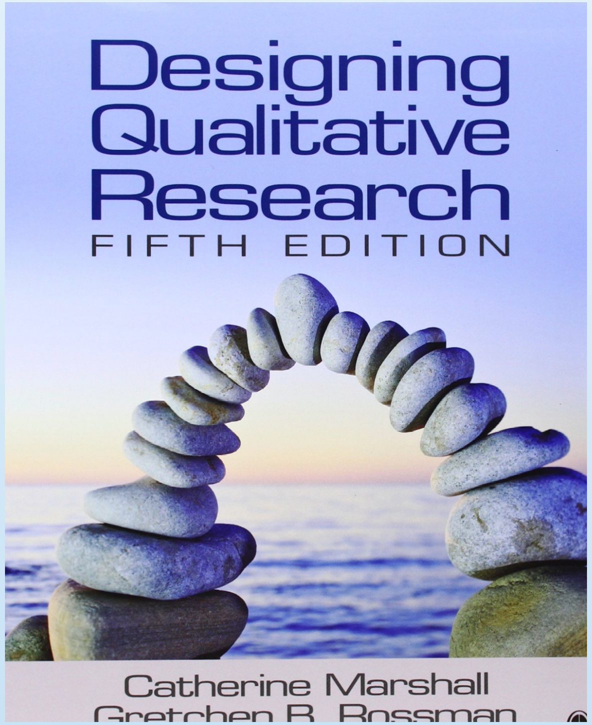
```
example (p q r: Prop) : ((p ∨ q) → r ) ↔ ((p → r)∧(q→ r)) :=
begin
split,
{intro h,
split,
{intro hp,
apply h,
left,
assumption},
{intro hq,
apply h,
right,
assumption}},
{sorry}
```

$$p \land (q \lor r) \leftrightarrow (p \land q) \lor (p \land r)$$

What I wish I knew when I started my first study

Quite unlike its pristine and logical presentation in journal articles real research is often confusing, messy, intensely frustrating, and fundamentally non-linear.

(Marshall & Rossman, 2014)



Marshall, C., & Rossman, G. B. (2014). *Designing qualitative research*. Sage publications.

In this talk

This is the story of two talks: one is a talk I usually give to mathematicians and mathematics educators when I present the work in Thoma and Iannone (2021) and Iannone and Thoma (2023) (slides with a white background) – the other is a talk about how the research happened (slides with a blue background).

International Journal of Research in Undergraduate Mathematics Education
https://doi.org/10.1007/s40753-021-00140-1

Check for updates

Learning about Proof with the Theorem Prover LEAN: the Abundant Numbers Task

Athina Thoma[1,3] · Paola Iannone[2]

Accepted: 24 March 2021 Published online: 07 July 2021

INTERNATIONAL JOURNAL OF MATHEMATICAL EDUCATION IN SCIENCE AND TECHNOLOGY
https://doi.org/10.1080/0020739X.2023.2178981

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Interactive theorem provers for university mathematics: an exploratory study of students' perceptions

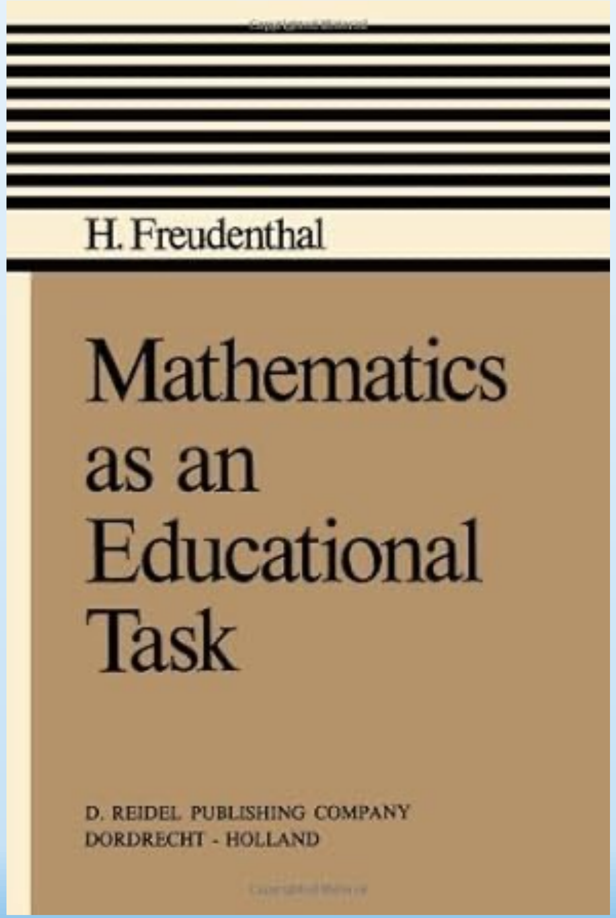
Paola Iannone[a] and Athina Thoma[b,c]

[a]Department of Mathematics Education, Loughborough University, Loughborough, UK; [b]Department of Mathematics, Imperial College London, London, UK; [c]Southampton Education School, University of Southampton, Southampton, UK

We often criticise the way in which mathematics is taught

The criticism concerns the fact that the way in which mathematics is researched is completely disconnected from the way in which mathematics is presented and taught.

However, we do the same in the social sciences!



Before I start

An apology to researchers using quantitative methods: some but not all the things I say will be relevant. Pre-registration of quantitative studies for example prevents cherry picking and *p*-hacking and allows for replicability – in this case research MUST happen the way in which it was planned.

BUT – I prevalently use qualitative methods – and here is where my expertise lies.

How it started

September 2018: a mathematician at Imperial College London plans to use an interactive theorem prover (ITP - Lean) to teach a first-year introduction to proof module – we had met at a meeting previously, so he contacted me to study the impact of using Lean on learning mathematics.

- I did not know what ITPs or Lean were at that time
- The course started a week after they contacted me – no time to plan!
- However, I was interested in this new technology as it concerned mathematical reasoning

Interactive theorem provers and the teaching of (pure) mathematics at university

Paola Iannone

School of Mathematics

University of Edinburgh

paola.iannone@ed.ac.uk


```
example (p q r: Prop) : ((p ∨ q) → r ) ↔ ((p → r)∧(q→ r)) :=
begin
split,
{intro h,
split,
{intro hp,
apply h,
left,
assumption},
{intro hq,
apply h,
right,
assumption}},
{sorry}
```

$$p \wedge (q \vee r) \leftrightarrow (p \wedge q) \vee (p \wedge r)$$

Programming in university mathematics

Mainly used for applied mathematics teaching (computational) and not pure mathematics.

Matlab, Mathematica, Maple, Python, R.

No evidence of the use of automated provers or programming for pure mathematics.

Sangwin and O’Toole (2017)

Proof at University – areas of difficulty

- Epistemological

  (Di Martino, Gregorio and Iannone, 2022)
- Proof comprehension

  (Mejia-Ramos et al., 2012)
- Proof writing

  (Lew and Mejia-Ramos, 2019)
- Proof appraisal

  (Selden and Selden, 2008)
- Proof composition

  (Weber and Alcock, 2004)

These difficulties have been documented in the literature – each may represent an area where impact of a pedagogical intervention may be found to be effective.

Proof at university: Proof comprehension and proof writing (Fukawa-Connelly, 2012)

- Definitions and their use
- Mathematical symbols and their use
- Logical status of statements and their links
- High level ideas
- Modular Structure of the Proof
- Use of examples

The gap between research and practice

- Programming has become an integral part of research in mathematics, but but not of mathematics teaching.

  Of the mathematicians surveyed 43% used computer programming in their research while only 18% included programming in their teaching. (Broley, 2016)

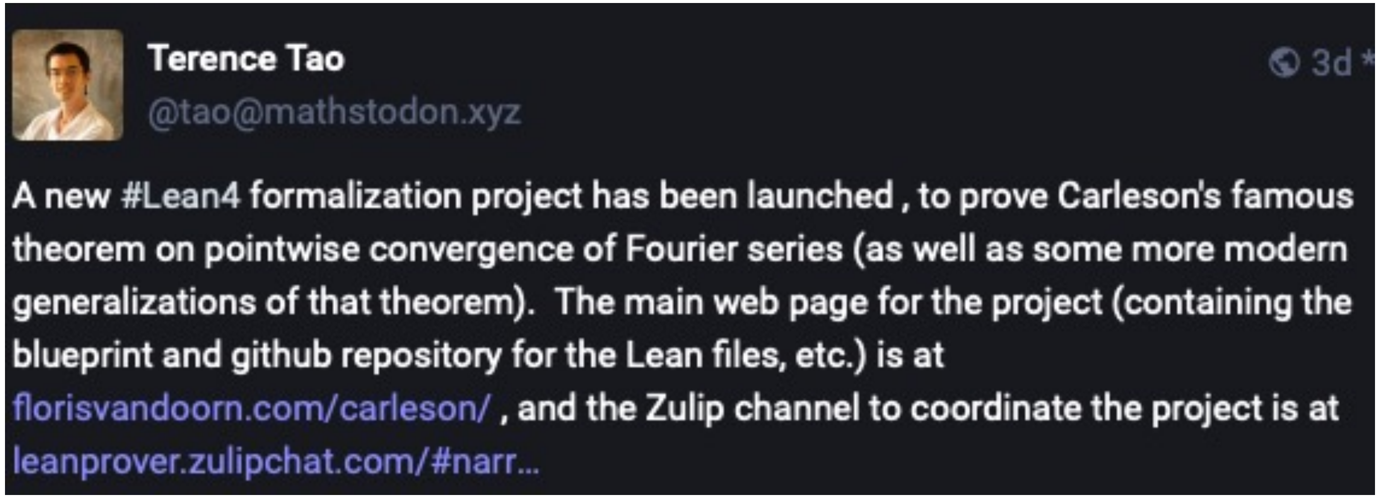
- Software use - aiming to assist students' proof production – rarely used.

Software designed by educators to assist students with pure mathematics and proof (e.g., ISETL https://www.swmath.org/software/1370_ - Dubinsky and Leron, 1994) is complete of documentation and manuals, has been shown to be effective – but is is hardly in use.

Interactive Theorem Provers

A proof assistant or interactive theorem prover **(ITP)** is a software tool to assist with the development of formal proofs by human-machine collaboration.

This involves some sort of interactive proof editor (programming interface), with a symbolic library of mathematical objects with which a human search for proofs, the details of which are stored in, and some steps provided by, a computer.



https://leanprover-community.github.io

The topic

One of the reasons why I took this assignment was the potential for Lean to bridge research and teaching – Artigue (2016) noted this pedagogical disconnect - with a pedagogical intervention. Perhaps following what Weber and Dawkins (2023) write about sustainability, this intervention may survive the initial enthusiasm of the lecturer who implemented it!

However, I had to quickly learn what are Interactive Theorem Provers (ITP) ☺ and Lean.

LEAN THEOREM PROVER

It brings interactive and automated
reasoning together
and build an interactive theorem prover with powerful
automation and an automated reasoning tool that produces detailed proofs.

LEAN has a rich language and can be used interactively.

Is built on a verified mathematical library and has as a programming environment in which one can:

compute with objects with precise formal semantics,
reason about the results of computation, and
write proof-producing automation.

https://github.com/leanprover/lean

THEOREM PROVER

Lemma

If x and y are natural numbers, and $y = x + 7$, then $2y = 2(x + 7)$.

```
lemma example2 (x y z : mynat) (h : y = x + 7) : 2 * y = 2 * (x + 7) :=
```

Proof :

```
begin
48
end
```

48:0: goal

```
x y z : mynat,
h : y = x + 7
⊢ 2 * y = 2 * (x + 7)
```

Lemma

If x and y are natural numbers, and $y = x + 7$, then $2y = 2(x + 7)$.

```
lemma example2 (x y z : mynat) (h : y = x + 7) : 2 * y = 2 * (x + 7) :=
```

Proof :

```
begin
48 rw h,
49 rw y
end
```

49:4: goal

```
x y z : mynat,
h : y = x + 7
⊢ 2 * (x + 7) = 2 * (x + 7)
```

49:4: tactic `rw`

49:0: error:

```
rewrite tactic failed, lemma is not an equality nor a iff
state:
x y z : mynat,
h : y = x + 7
⊢ 2 * (x + 7) = 2 * (x + 7)
```

Frustrations about researching technology

One of the frustration about researching the use of technology in teaching/learning mathematics is that technology moves much faster than I can analyse data!

This study used version 3 of Lean – version 4 is now out with new and better features.

Therefore, we need to be conscious about features of the tool and how these may shape findings which may be version dependent.

Framing the research

The study was planned to be about the impact that using an ITP had on students' understanding of proof. This is vague so we needed a better frame to describe precisely what we wanted to investigate. To this aim we deployed what we knew about difficulties with proof in manageable categories.

This as we will see will allow us to frame our results.

The plan

Lean was not a compulsory part of the course – the course could be taken without engaging with Lean. We hoped that a good number of students would take it up – and the we could for example make inferences as to whether engaging with Lean helped the students being more successful at writing correct proofs.

Previous research had hinted that this may be the case (e.g., Avigad, 2919).

Disaster strikes

Through an initial questionnaire we realised that only 18 out of 300 students were still working with Lean in week 3 of the course!

We had to change both the methods we had planned and the research questions we wanted to ask – the number of students engaging with Lean was too small to tell us anything about attainment in proof production.

The research questions

RQ1: What characteristics are observed to be common to proofs by students who engaged with the software Lean?

RQ2: What are the barriers students encountered during their engagement with Lean?

The second question is about students' perceptions and engagement – we will not talk about this in this talk.

Context

- A Year 1 course: Transition to proof
- Hourly lectures (3 per week), weekly seminar classes, Lean **not** compulsory
- And optional Lean sessions on Thursday evenings

300 students in the course – large research-intensive university in the UK.

Data collection

Data	When	Purpose
Questionnaires	October-December	Students' engagement with LEAN
Interviews	November/ December	Students' proof writing
Marks for module tests	October-December	Assessment
Marks for the final exam	January	Assessment

To ascertain the uptake of LEAN workshops – 18 students (out of 300) attended the workshops - and to investigate students' perceptions of the use of this programming language as part of their instruction. December questionnaire N=99.

Interviews with 37 volunteers lasted about 1 hour. Various tasks were given with familiar and unfamiliar mathematical objects. We focus on the abundant number task. The proofs were scored by a researcher outside the project.

To ascertain where there was an attainment difference between the students who took up LEAN and those who didn't at the start and the end of the teaching period. There was no difference in attainment at the start of the term – but there was a difference (expected) at the end

The new focus and the corresponding new methods – RQ1

The focus is now on the transfer of desirable proof habits from Lean to writing proofs with pen and paper. To this aim then task-based interviews seemed to be the best tool.

We needed a framework that would help us making the proofs produced in the interviews manageable - this is why we used the Selden and Selden (2007) distinction in formal –rhetorical and problem – centred parts of the proof. Once we had this way of separating the two parts of the proof, we looked for a framework for proof comprehension and proof writing – we adopted the one from Fukawa-Connelly (2012).

One more issue

The students who took up Lean spent more time on average ‘doing mathematics’, were probably more motivated and keener. We needed to make sure that they were not a self-selecting sample of students who were higher achiever in their class at the start of the course.

To find this out we analysed the class results from a test taken at the start of the course – these showed no significant difference of achievement.

Lastly – to compare proofs at the same achievement level we asked a mathematician external to the project to score the proofs written during the interviews – for one task.

The task

Definition: An **abundant number** is an integer n whose divisors add up to more than $2n$.
Definition: An **perfect number** is an integer n whose divisors add up to exactly $2n$.

Task: If n is perfect, then kn is abundant for every $k \in \mathbb{N}$. Investigate.

Discussion: The statement is false for $k = 1$.
Let $k \in \mathbb{N}$ with $k > 1$ and let n be a perfect positive integer.
Let $d_1, d_2, \ldots, d_r$ be the divisors of n. We have $\sum_{i=1}^{r} d_i = 2n$.
Consider now $kd_1, kd_2, \ldots, kd_r$, these are among the divisors of kn and we have that $kd_i \neq 1$ for each i as $k > 1$.
1 is also a divisors of kn. Therefore kn has amongst its divisors kd_1, kd_2, $\ldots$, kd_r and 1 their sum is $2kn + 1$ which is greater than $2kn$.
This implies that kn is abundant.

Formal Rhetorical part

The task and a correct solution- With parts of a proof*

The problem centred part

*From Selden and Selden, 2008

The task

The choice of task for the interviews is crucial – the task needs to be complex enough to allow for some real mathematical thinking but not too difficult that no student will be able to tackle it.

In our case we selected an unseen task – students were not aware of the definition of abundant number and would have not seen the proof we asked them to engage with. We wanted to avoid memory recall which may come from seen proofs (Azrou & Khelladi, 2019).

Data Analysis

Phase 1

Zazkis, Weber and Mejia-Ramos (2015) scoring scheme for abundant number task in the interviews.

Proofs were scored from 4 (complete and correct proof) to 0 (no attempt was made) – analysis in phase 2 carried out on proofs at the same scoring level. Scoring done by a mathematician external to the project.

Preparatory phase for Phase 2.

Phase 2

Fukawa-Connelly (2012) theoretical framework – analysis of the proof outputs on proofs at the same scoring level:

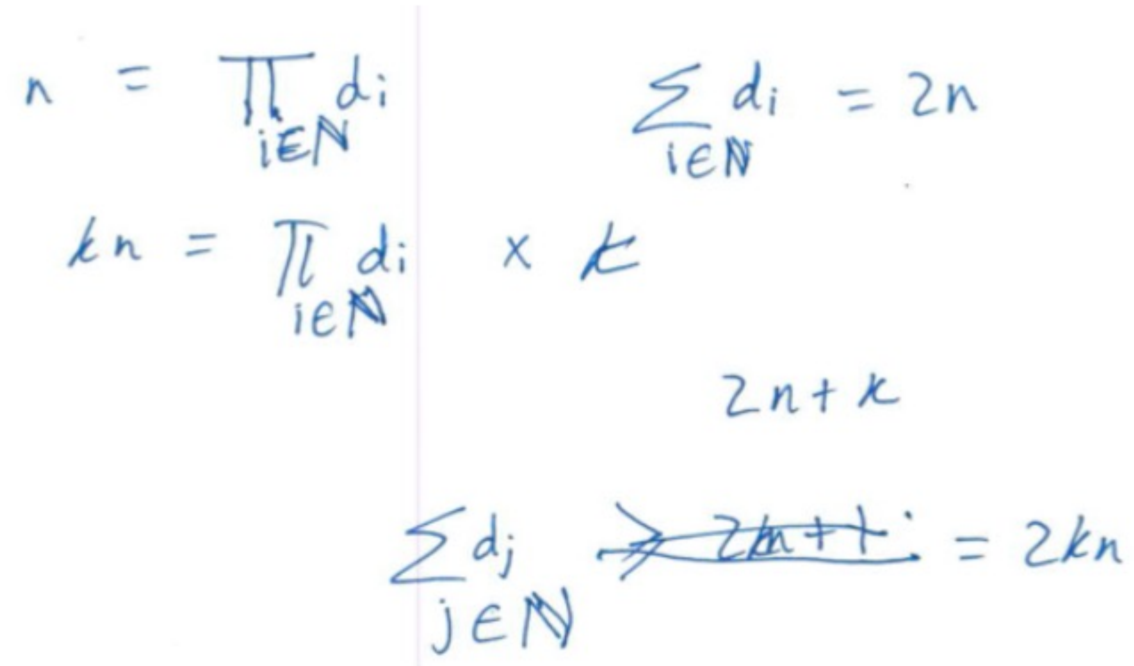
- Definitions and their use
- Mathematical symbols and their use
- Logical status of statements and their links
- High level ideas
- Modular Structure of the Proof
- Use of examples

The theoretical framework

We operationalised the theoretical framework with the following coding schemes:

Definition Given [DEF-GIV] Using familiar definitions [DEF-FAM] Using familiar statements/terms [STAT-FAM] Introduction of symbolism [SYM] Logical status of statements [LOG-STAT]	Links to previous statement [LOG-FOL] High-Level ideas [APP] Modular structure of the proof [FIT] Example Use [EX]

With an example of analysis



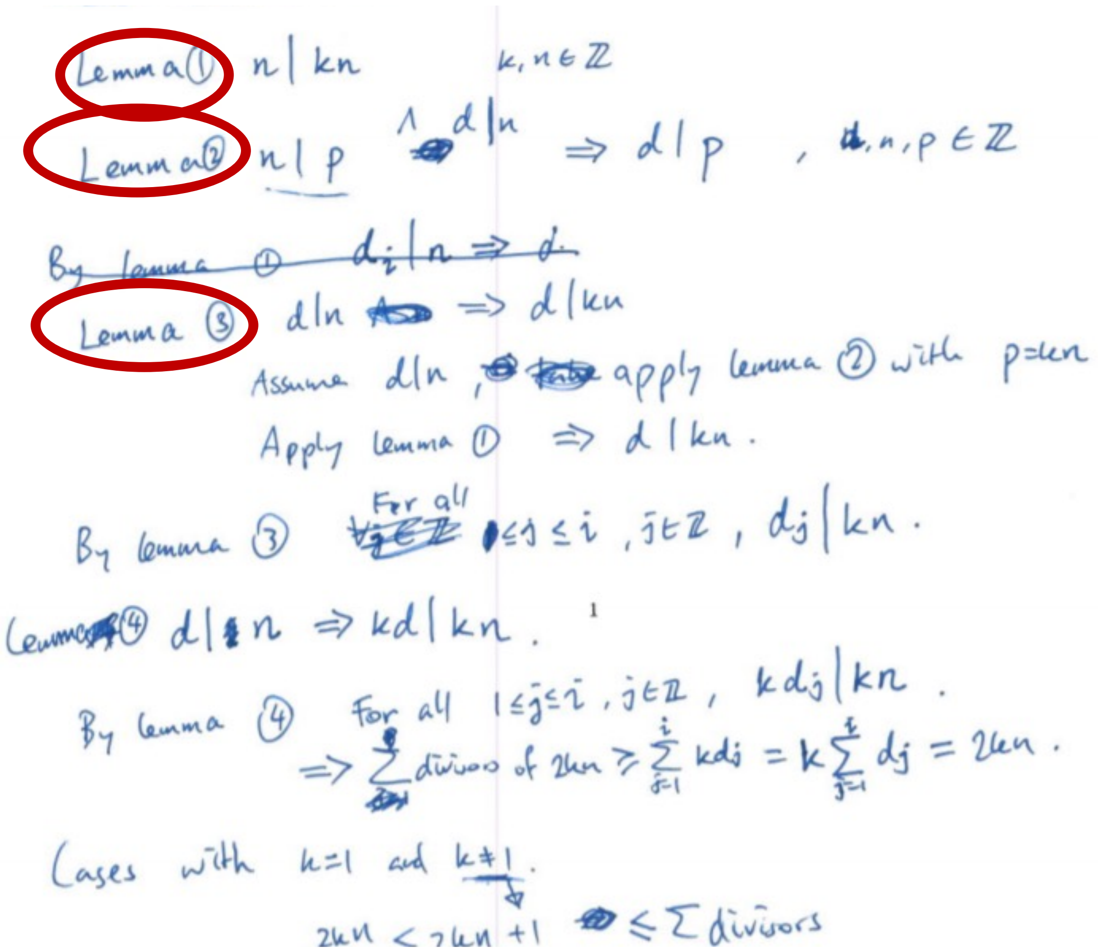
[1] I don't have any leads on the first part but I will try the first part to start with
[2] Let's try n equals the product of d - for divisor - d_i, i equals. [DEF]
[3] Let's say i integer I mean natural numbers. Then kn is equal to the product of the divisors times k. [LOG FOL]
[4] And because by assuming that n is perfect here. We say the sum of factors/divisors are equal to $2n$ [LOG FOL]

Data Analysis

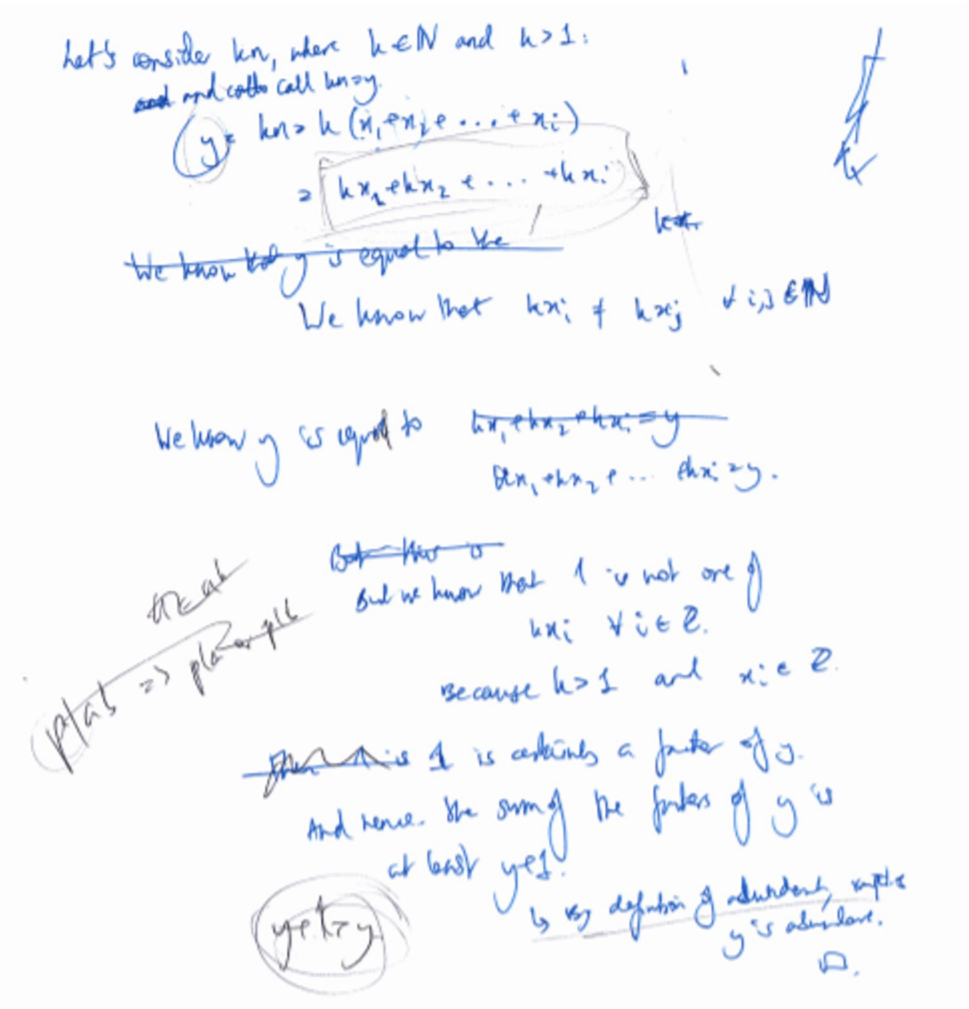
In the full papers we show how the analysis of the data deploys the Fukawa-Connelly (2012) framework and the Selden and Selden (2007) proof structure. Here because of the mixed audience for which the talk was designed, I present only some results.

It is important: once a theoretical framework is chosen it must be linked both to the data analysis and to the inferences drawn from the data.

Score 4: Second part of the proof

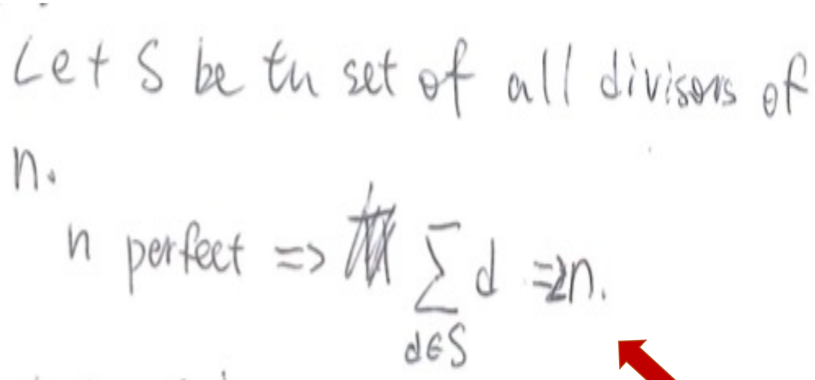


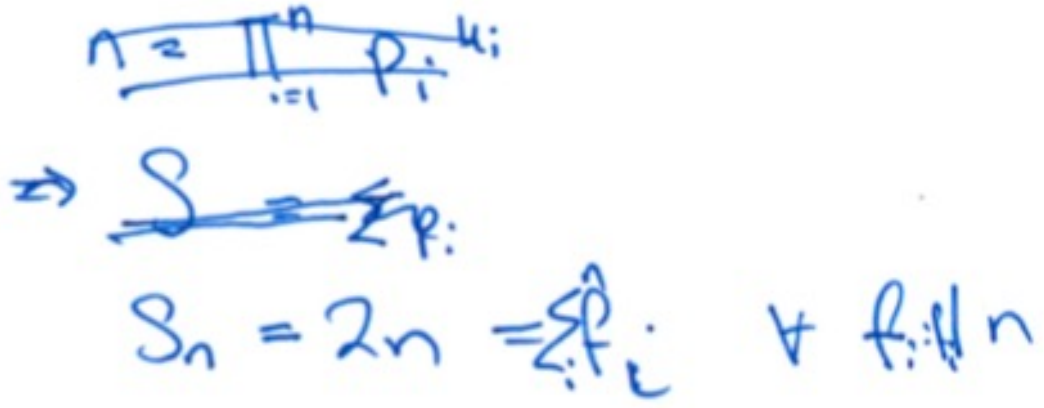
Lean user



Not a Lean user

Score 3: Start of the proof



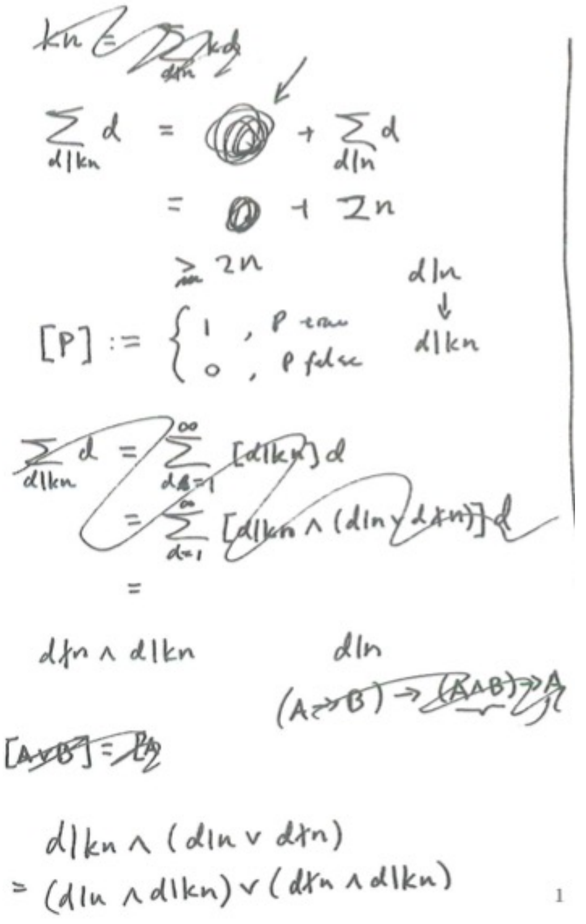


Use of mathematical language

Lean user

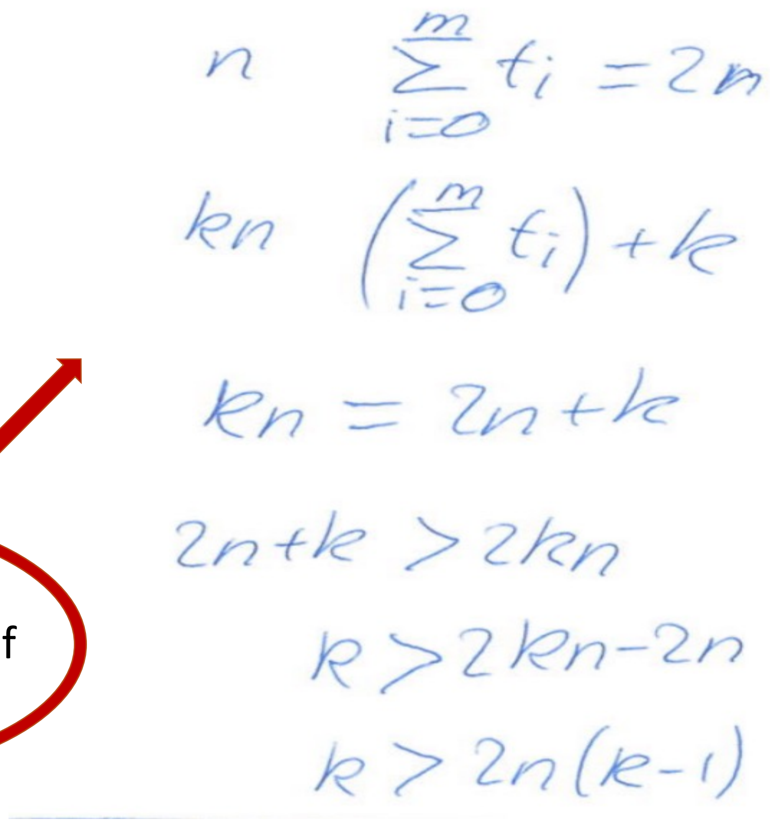
Not a Lean user

Score 1: Second part of the proof



Attempt to formalize when not needed – Role of intuition?

Confusion and Inaccurate chain of deductions



Lean user

Not a Lean user

One more note on the framework we used

Even in this quick presentation we see how thinking of proofs as in Selden and Selden (2007) allowed us to see characteristics of the two parts while the Fukawa-Connelly (2012) one allowed us to focus on language, mathematical writing etc.

Some results

Mathematical writing:

- Use of technical mathematical language and symbolism.
- Explicit statements regarding where certain mathematical objects belonged.
- Precise introduction of the mathematical objects that play a role in the proof.
- Use of words and punctuation to accompany the mathematical symbols.

Proof structure:

The breakdown of the proof goal in intermediate sub-goals (often overt)

Some remarks

This is a small exploratory study, but it suggests that:

- There can be a transfer between mathematical habits acquired while learning Lean to writing proofs on pen and paper
- The distinction between natural language, technical language (of mathematics) and programming language becomes clear to the students
- The emphasis on formalisation may help students transition to more advanced epistemologies at the start of their degree.

The message

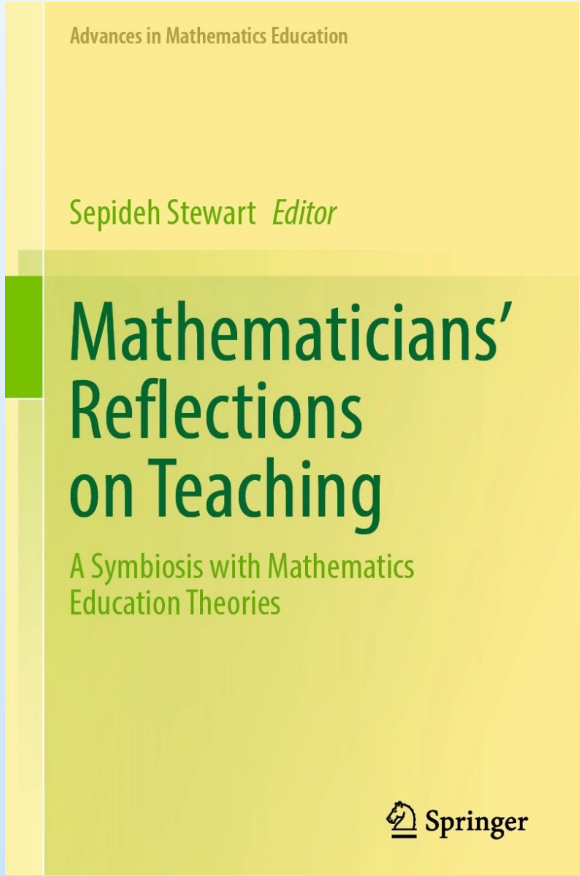
As researchers the message we got from the study is that the intervention as it stood was not successful – not many students engaged with Lean and those who did were the motivated ones (they have in fact all finished PhDs in mathematics or computer science by now!).

How do we convey the message to the very enthusiastic lecturer and outstanding mathematics researcher who allowed us in their classroom?

Collaborative research

I have often written about collaborative research* - and much of my research has been collaborative with mathematicians. Indeed, this is the way I started to do educational research, by collaborating with a mathematics educator.

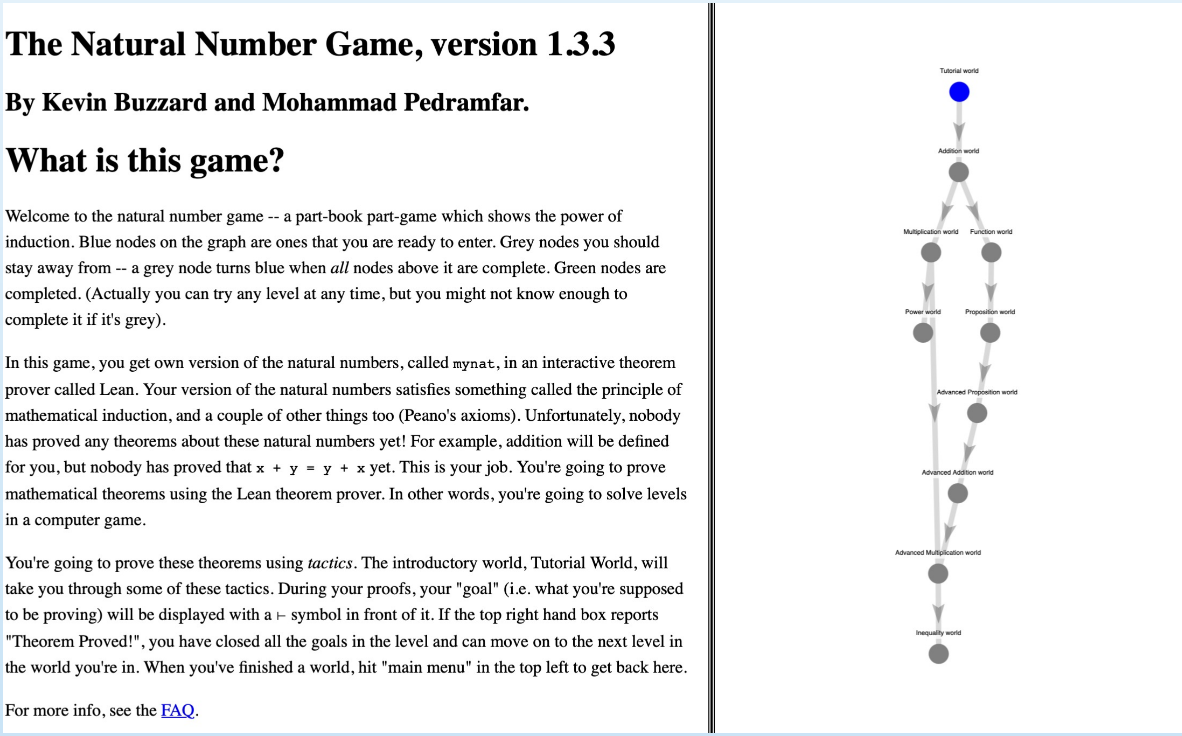
In Iannone (2023) I make the case for collaborative research for the (realistic) evaluation of teaching interventions.



Our collaborative work

We were lucky – the lecturer we worked with took our suggestions on board and reflected on our results – in fact spoke about this work at relevant conferences. The outcome of our conversation was the Natural Number Game.

It is very important for this kind of research to maintain the collaborative aspect.



The Natural Number Game, version 1.3.3

By Kevin Buzzard and Mohammad Pedramfar.

What is this game?

Welcome to the natural number game -- a part-book part-game which shows the power of induction. Blue nodes on the graph are ones that you are ready to enter. Grey nodes you should stay away from -- a grey node turns blue when *all* nodes above it are complete. Green nodes are completed. (Actually you can try any level at any time, but you might not know enough to complete it if it's grey).

In this game, you get own version of the natural numbers, called `mynat`, in an interactive theorem prover called Lean. Your version of the natural numbers satisfies something called the principle of mathematical induction, and a couple of other things too (Peano's axioms). Unfortunately, nobody has proved any theorems about these natural numbers yet! For example, addition will be defined for you, but nobody has proved that `x + y = y + x` yet. This is your job. You're going to prove mathematical theorems using the Lean theorem prover. In other words, you're going to solve levels in a computer game.

You're going to prove these theorems using *tactics*. The introductory world, Tutorial World, will take you through some of these tactics. During your proofs, your "goal" (i.e. what you're supposed to be proving) will be displayed with a ⊢ symbol in front of it. If the top right hand box reports "Theorem Proved!", you have closed all the goals in the level and can move on to the next level in the world you're in. When you've finished a world, hit "main menu" in the top left to get back here.

For more info, see the FAQ.

Some remarks

Teaching mathematicians to use computer theorem provers

Kevin Buzzard

Teaching with ITPs

Theorem (Iannone, Thoma)

Asking a student to learn new mathematics and at the same time to learn how to use an interactive theorem prover, is asking too much.

5

From Kevin Buzzard's talk at the first Teaching with Lean conference

- Trying to teach both complex mathematics and complex programming at the same time in the first year will not be successful for most students
- We do not know how students use Lean (or any ITP) as a toll
- What is the role of intuition when programming an ITP?

Some final remarks

I hope this was a useful example of doing research related to the use of technology in university mathematics (sorry Olov (university mathematics) and Andreas (proof and argumentation) for stealing your topics ☺).

Above all I hope I have conveyed the messiness of the process of research and how decisions can be taken to alleviate the messiness and carry out research as a systematic enquiry.

How it is going

- Our research collaboration with mathematicians on the use of Lean is continuing
- The Lean community is growing internationally, educational research concerning ITPs is also gaining momentum
- ITPs may be the most important development in mathematics of recent years which may change mathematicians’ practices

Thank you for listening!

If P, Q and R are true/false statements, then P $\iff$ Q and Q $\iff$ R together imply P $\iff$ R.

Proof :

```
begin
  intro h,
  cases h with pq qp,
  intro j,
  cases j with qr rq,
  split,
  intro p,
  apply qr,
  apply pq,
  exact p,
  intro r,
  apply qp,
  apply rq,
  exact r,
end
```

Proof :

```
begin
  intros h1 h2,
  cases h1 with h3 h4,
  cases h2 with h5 h6,
  split,
  repeat {cc},
end
```

THE UNIVERSITY *of* EDINBURGH