Faculty of Environment, Science and Economy

MTH3035 Mathematics Group Project

Group 9B: Improving feedback for student proofs with Lean and AI

Academic Year 2023 / 2024

Candidate Numbers: 214389, 187357, 145764, 109519, 200844, 104408, 183451, 167904

Contents

1 Abstract

This research project is centred around improving feedback for students' proofs using Lean and AI, aiming to assess Lean and AI's effectiveness in improving students' comprehension of mathematical proofs. The significance of this investigation lies in the foundational role that proofs play in advanced mathematics education, with their comprehension being a crucial aspect of academic success.

The research explores if there is an interface between Lean and generative artificial intelligence (AI) to aid students in mathematical proof construction and comprehension. The question at the core of our investigation is: how effective are these technological interventions in providing personalised feedback and support to students struggling with individualised proof questions?

Our research aims to assess the effectiveness of Lean and AI in assisting students with mathematical proof comprehension. Building on this assessment, we intend to develop additional tools, namely a website and a handout, that aid students in their study of proofs, emphasising a deeper conceptual understanding and effective response to feedback from technological interfaces.

Our approach involves a comprehensive evaluation including: focus groups, surveys, interviews with educators, and an exploration of the limitations and capabilities of AI. This multifaceted methodology allows us to gather nuanced insights into the impact of technology on students' proof comprehension.

Our research highlights the potential of integrating advanced technological tools to bridge gaps in mathematical education, offering personalised support to students and building a strong understanding of mathematical proofs. The research provides insight into enhancing students' proficiency in mathematical proof comprehension through improved feedback.

Our study revealed that effectively integrating Lean and AI into learning enhances students' confidence and proficiency in mathematical proof construction. Yet we struggled to find a direct interface between Lean and AI. This was primarily the result of AI's lack of knowledge of Lean. We managed to utilise them both in our website; using Lean to help construct proof and AI prompts to increase students' comprehension of the theory behind mathematical proofs. Via extensive feedback and evaluation of such, we discovered our additional tool helped improve students' comprehension of mathematical proofs, and the tools would therefore be recommended to their peers or colleagues.

Therefore, our findings emphasise the promising potential of advanced technologies in transforming mathematical education, offering tailored assistance, and significantly improving students' proficiency in navigating proof complexities. Additionally, our study shows a continuous literature review and research of pedagogical advances in this field.

2 Introduction

2.1 Mathematical Proofs

2.1.1 The background and importance of mathematical proof

As mathematics students make the transition into higher education, they are introduced to various new concepts, among which, mathematical proofs commonly stand as a significant challenge. In mathematics, where logic and reasoning are crucial, proofs play a central role in confirming the validity of conjectures and theorems. However, this concept presents considerable challenges for new learners. The rigorous nature of proofs and proof-based learning often pose difficulties for students, making it a critical area for educational intervention [Web08]. Because of these challenges, many students regard proofs and learning proofs as unimportant or useless [HM22].

Proofs originate from a committed effort to strictly adhere to logical principles and apply formal rules of inference with diligence. These proofs are built upon deductive reasoning; they stick to a structured and systematic process where conclusions are derived from established premises through logical reasoning. This form of deductive reasoning originates from the writings of Aristotle and was evolved by the likes of Euclid, who laid the groundwork for formal proof systems. This led to mathematicians and philosophers such as René Descartes improving upon the systems already in place by introducing the "Method of Doubt" in his book [Des41], which is a methodical approach to questioning the truth of one's belief. This emphasises the importance of clear and certain reasoning. This is further emphasised by his book [Des37], which introduced rules to follow when proving an idea deductively.

2.1.2 What it is to understand a mathematical proof and how can we evaluate this?

Dr Gihan Marasingha, a lecturer from the University of Exeter, defines comprehension of mathematical proofs as "constructing reasoned arguments from axioms or theorems through step-by-step logic to arrive at a conclusion". Understanding proofs involves composing and comprehending these logical arguments, forming a fundamental pillar in advanced mathematics education [All97]. However, students often encounter challenges in grasping and constructing proofs.

Students commonly face a diverse range of challenges when they encounter mathematical proofs. This has resulted in a high demand for more personalised support, often in the form of an individualised feedback mechanism. Furthermore, we discovered discrepancies between mathematicians' practices and educational approaches contribute to students' disinterest and lack of appreciation for proofs [HM22]. Also, evaluating the validity of proofs poses as a significant hurdle, revealing a lack of understanding of what constitutes as a solid argument. This highlights the necessity of explicitly guiding students in developing their critical thinking and logical reasoning skills as they are essential for comprehending and constructing rigorous mathematical proofs [HM22].

Understanding the dimensions integral to grasping mathematical proofs involves assessing several components. Conceptual grasp evaluates the clarity of foundational concepts like definitions and axioms involved in proof construction [HGS06]. Logical coherence scrutinises the progression of ideas, the validity of logical deductions, and the avoidance of fallacies[MRI09]. Transferable skills encompass the ability to apply proof strategies across various mathematical scenarios. Problem-solving proficiency aligns with understanding proofs, requiring the breakdown of complex problems, the application of appropriate techniques, and the construction of coherent arguments[HGS06].

Moreover, an effective way to evaluate students' comprehension of a mathematical proof is to observe their ability to construct their own proofs. When students can articulate and explain a proof to their peers, walking through each step of their reasoning, it serves as a demonstration of their understanding. This process not only gauges their grasp of the specific proof in question but also reveals their broader capacity to apply logical principles of inference in a coherent manner [MRFW+12a]. The ability to communicate mathematical ideas demonstrates a deeper understanding, as it requires the integration of conceptual understanding with effective verbal expression, resulting in a more comprehensive evaluation of their overall proficiency in understanding a mathematical proof.

2.1.3 Teaching proofs and its various teaching methods

Effective proof teaching presents unique challenges in the ever-changing field of mathematics education. Researchers explore students' preferences in teaching methods, shedding light on the effectiveness of both traditional and technology-based approaches.

According to [CG22] many students prefer more active methods of teaching where they are encouraged to actively participate in lessons by thinking, researching and discussing. Students practise skills, work through issues, grapple with challenging questions, come to decisions, and offer answers to articulate concepts in their own terms. One example of an active teaching method that students highly prefer is “guided discovery” which entails educators facilitating learning through carefully designed activities that guide students to discover key concepts and principles independently. This method places an emphasis on active engagement, critical thinking, and problem-solving, allowing students to explore and make connections independently with the guidance of the teacher.

In a proof learning setting, active teaching methods encourage students to actively engage with the process of proving mathematical theorems. Instead of presenting proofs as static information, students are prompted to explore and discover the underlying logic and structure. Furthermore, this provides a structured environment for students to explore various proof techniques. Through carefully designed activities, students can discover the methods and strategies employed in constructing mathematical proofs. Moreover, learning about proofs involves building logical connections between statements and understanding the reasoning behind each step. Additionally, as shown in [SFW14], active teaching methods allow students to make these connections on their own; or with the help of their peers in the case of "cooperative learning" which is an instructional strategy where students work together in small, diverse groups to achieve common learning goals, fostering a deeper comprehension of the logical structure inherent in mathematical proofs.

2.1.4 Impact of Web-based resources

The integration of computer programming and computer-assisted instruction (CAI) in mathematics education holds significant promise, as highlighted by several studies. There is a substantial disparity between professional mathematical practices and educational curricula, emphasising the need to bridge this gap by integrating programming into teaching [Bro16]. Forsström and Kaufmann echo this sentiment, underscoring the benefits of programming in enhancing student engagement [KS21], attitudes, and real-world applications of mathematical concepts [FK18]. De Witte and Soliman's research collaborates the effectiveness of CAI programs in improving learning outcomes, demonstrating higher test scores and enhanced comprehension among students engaged with such tools [SH16]. Overall, these findings emphasise the potential of programming and CAI to bolster understanding, improve learning outcomes, and foster positive attitudes toward mathematics. However, the integration of these tools requires clear guidelines and a concerted effort to align professional practices with educational curricula for the betterment of mathematics education.

2.1.5 Lean principles in education

In our pursuit to enhance feedback for student proofs, the convergence of Lean (an interactive and automated theorem prover) and AI tools offers a promising avenue. Lean, rooted in precise mathematical representation, aids in defining, reasoning, and computing various mathematical structures, crucial for teaching and learning mathematics. It fosters rigorous proof development and automation, supporting students in complex mathematical reasoning [EUR+17].

However, student studies reveal challenges in programming and emotional hurdles with Lean, despite its positive impact on structured proof-writing approaches and perceptions of rigorous proofs [IT23]. Engagement with Lean software showcases benefits in precision, communication, and structured proof writing, albeit requiring further exploration of reasoning approaches [AP21].

Despite challenges, Lean's potential in mathematics education is significant, bridging interactive and automated theorem proving while remaining accessible through its open-source nature [EUR+17]. Our project aims to leverage these insights by developing a website [FK18] that combines Lean and AI prompts. By understanding Lean's principles, acknowledging its challenges, and leveraging its potential in mathematics education, our goal is to create an innovative tool optimising, feedback for students' proof-writing endeavours [IT23].

2.1.6 AI in mathematical education

Insights from studies on generative AI models, particularly ChatGPT, offer valuable perspectives into ways to improve feedback for students [Meg23]. ChatGPT showcases promise in transforming practices across Statistical Process Control (SPC) and education, serving as an effective idea-generation tool. However, while excelling in some structured tasks like error message translations and explaining familiar concepts, it faces challenges in handling specific Lean questions, as it is not as well documented as other programming languages such as Python. Therefore regarding specific Lean questions, ChatGPT often yields misleading results, calling for additional supplementary methods [Meg23]. Further ChatGPT is more reliable for the theory behind mathematical proofs rather than answering questions regarding Lean.

Analysis of ChatGPT's mathematical capabilities demonstrates distinct proficiency levels across difficulty levels and data sets [Fri23]. While it proves useful for fact querying and basic knowledge interfaces, it demonstrates limitations in tackling graduate-level complexity. Contrary to media portrayals, its overall performance in mathematics remains below the graduate student level, cautioning against relying solely on ChatGPT for advanced mathematics exams [Fri23].

Exploration of ChatGPT's role across various domains highlights its potential benefits in enhancing productivity and problem-solving but also emphasises risks tied to over reliance [AAR23]. These risks include erroneous responses, code inaccuracies, limited logical reasoning, and ethical concerns. To mitigate these risks, strategic methodologies and independent verification of outputs are proposed, emphasising the need for expert guidance and iterative interaction with ChatGPT [AAR23].

However, it is crucial to acknowledge the rapid evolution of generative AI. While current studies offer valuable insights, the swift advancements in AI suggest that research may quickly become outdated. Hence, while informative, understanding the evolving nature of generative AI remains essential when utilising these tools effectively for improving student proof.

2.2 Our Aims and Objectives

In our collaborative project, defining clear aims and objectives played a pivotal role in shaping our research trajectory and ensuring focused efforts. Dissecting each aim and objective was vital for clarity and unanimity within the group, ensuring alignment on our intended steps and goals. Our overarching aim was to enhance students' comprehension of mathematical proofs through the evaluation of innovative tools like Lean and artificial intelligence (AI). To achieve this, we established specific objectives:

2.2.1 Assessing the Effectiveness of Lean and AI in Assisting Students with Mathematical Proof

The primary objective of our project, and thereby this report intends to critically assess the effectiveness of the mathematical theorem prover Lean, which aims to verify and produce detailed proofs [AP21], as well as generative artificial intelligence (AI), such as ChatGPT, on aiding students with mathematical proof. The focus revolves around evaluating the feedback mechanisms inherent in these interfaces, particularly concerning their capacity to guide students through proof questions. Via empirical investigation and rigorous analysis, we aim to discuss the strengths and limitations of these technological tools as educational aids in student's comprehension of mathematical proofs to assist us in producing a reliable resource and to work out the most beneficial format for students.

2.2.2 Evaluate the Impact of Lean on Student Proof Comprehension

We aim to conduct a thorough investigation to understand how Lean affects students' capacity to create and understand mathematical proofs. This involves looking closely at how the tool is set up, how it works, and how it influences the way students grasp the structures of mathematical proofs. In simpler terms, we want to figure out how Lean, like any software, helps or hinders students when they're working on mathematical proofs, considering questions like how accessible it is and how it shapes their understanding of proof concepts. We aim to do this so we can gain critical information about the best aspects of Lean to assist us in producing similar aspects to incorporate into our educational tools.

2.2.3 Assess the Role of Generative AI in Enhancing Student Feedback

Further, we will explore the contributions of generative AI, specifically ChatGPT, in providing meaningful and accurate feedback to students engaged in mathematical proof activities. ChatGPT is a language processing tool driven by AI, allowing for human-like conversation, it is commonly utilised by students to compose emails, essays and code. [Ort23]. This objective aims to investigate AI's ability to offer personalised guidance, identify common errors, and facilitate a more nuanced as well as imminent understanding of proof strategies. These pieces of information will be useful going forward because it is necessary to be aware of AI's potential in the teaching of mathematical proofs so we can integrate it into our tools appropriately.

2.2.4 Development of an additional tool for Students Studying Mathematical Proof

To extend upon the aims listed above, we will strive to leverage the knowledge acquired to create a practical tool that offers additional assistance to students immersed in the study of mathematical proofs, more specifically the study of the First Year Mathematics module, MTH1001. We chose to focus on this group of students because they work very closely with the Lean theorem prover to assist them in their acquisition of mathematical proof knowledge. The tool is designed to deepen students' comprehension of proof concepts as well as improve the feedback they receive by utilising the interface between Lean and AI. This tool will be easy for users to navigate, combining the features of Lean and AI to offer students a smooth and user-friendly learning experience. We aspire to create a website, as digital platforms have become vital in the landscape of modern education [FK18], especially after the COVID-19 Pandemic forced education to become more flexible [TRI22]. The pandemic resulted in students becoming more comfortable with online learning resources and we therefore believe students will appreciate a website and find it accessible. This goal underscores the significance of building a platform that is easy to access, straightforward to use, and helps students better grasp mathematical proofs. In simpler terms, we want to make a tool that brings together Lean and AI in a way that's easy and intuitive for students, enhancing their understanding of mathematical proof.

2.2.5 Facilitate Adaptive Learning Approaches Through Feedback Response

A major emphasis of our research was on assessing the feedback given from such technologies, as we wanted to ensure that with the implementation of our resources students understood the provided feedback and found it constructive. Hence we aimed to create an approach within the tools that guides students via effectively responding to feedback from both AI and Lean. Feedback is proven to be vital in the learning process to improve a students performance [oSC23]. In order to be effective, feedback requires an explanation of what they are doing correct and incorrectly, it should be given in a timely manner and should be individualised to meet the needs of the student [oSC23]. Utilising Lean and AI, we aim to create effective feedback responses with instant and personalised feedback for each student. This is a progression from traditional teaching feedback which is not usually received instantly and also an improvement from general mark schemes which are not unique to each student. This objective involves developing strategies to help students interpret and incorporate feedback into their proof-solving processes, ultimately fostering a more personalised learning experience.

2.2.6 Align tool design with Educational Literature and Pedagogical Advancements

Further, with the ongoing nature of our literature review, we will ensure that the developed tools align with the latest educational literature and pedagogical advancements in the field of mathematical education. We are aware that staying current in our field of research will ensure our study 'fits into the larger context of scientific knowledge and prevent duplicating work that's already been completed'. We will also rely on other researchers to understand specific elements related to this discipline [Man22]. This objective necessitates an ongoing commitment to staying up to date with the latest educational practices to optimise the effectiveness of our tools.

2.2.7 Enhancing Overall Understanding of Mathematical Proof

Throughout our aims and objectives, we are dedicated to contributing to a broader understanding of mathematical proof among students. This will be achieved by combining the insights gained from assessing Lean and AI; developing supportive tools and aligning with educational literature. The overarching goal is to help students better comprehend mathematical proofs and learn about them more effectively. This holistic approach aims to bridge the gap between theoretical knowledge and practical application, as well as improve the feedback they receive via Lean and AI. In turn, it equips students with the skills needed to navigate learning and understanding mathematical proofs.

3 Methodology

3.1 Project Overview

3.1.1 Planning and Risk Analysis

When starting our project it was important for us to plan out how we would achieve our end goals [Arm82]. By creating milestones, depicted in our Gantt chart we had specific objectives to achieve in a time frame, this was to ensure each aspect of the project was completed with sufficient time. Strategic planning is a vital part of any project and should be done thoroughly to achieve the greatest results [Ste10]. In order to ensure we kept to our plan and everyone had a role, we first made a table which included what was to be done, who was doing it and the date it had to be completed. This helped us to evenly distribute the tasks to guarantee every member was doing their fair share of the workload. Additionally, this gave us deadlines to work towards to ensure we were consistently keeping on top of the tasks. This plan was used in the creation of our Gantt Chart which is demonstrated by Figure (2). An example from our project plan can be seen below in Figure(1).

09/11/2023	Update Presentation Construction	Kofi, James, Rhianna, Amelia
14/11/2023	Update Presentation Preparation	Kofi, James, Rhianna, Amelia
24/11/2023	Final Presentation Construction	All
01/12/2023	Final Presentation Preparation	All
06/11/2023	Report Writing	Sam, James, Amelia, Rhianna
06/11/2023	Structure and Introduction	Amy, Jessica
13/11/2023	Methodology	Jessica
27/11/2023	Findings	Shohid, Jessica
27/11/2023	Conclusion and alterations	All
20/11/2023	Evaluation	Sam, James, Amelia, Rhianna
20/11/2023	Comparison of Solutions	Amy
28/11/2023	Exploration of Additional Solutions	All
27/11/2023	Provide Feedback to the University	All
04/12/2023	Focus Group 7	Amy, Shohid, Sam, Jessica

Figure 1: Example of our plan table

To ensure fairness, tasks were distributed among group members based on their individual strengths. This strategic allocation allowed us to construct a Gantt chart, illustrating the critical path we would follow. Refer to Figure(2) for a visual representation of our project timeline.

Figure 2: Gantt Chart showing critical path

The chart shows a sequence of tasks with dependencies, highlighting the necessity for certain items to be completed before others begin. Additionally, certain tasks span multiple weeks and therefore were started as soon as possible to optimise time efficiently. We opted for a Gantt chart due to its simplicity and effectiveness in visualising task timelines. Its straightforward format allows for clear comprehension of what needs to be accomplished and when, making it a practical tool for our project management [Kum05].

When addressing risk analysis, we prioritised anticipating potential challenges in order to proactively develop preemptive solutions, thereby minimising disruptions to our project. This comprehensive approach ensures that all risks are thoroughly considered, and the group understands the level of threat posed by each identified risk based on assigned numerical values [Ave12]. We have produced a risk analysis table with 5 columns: the risk, the likelihood of the risk (L) between 1-5, the impact of the risk (I) between 1-5, the total (L x I) and finally, what can be done to mitigate the risk. Creating a total column by multiplying the two factors, gave us an easy visualisation of what risks would be the most prominent and the ones to prioritise. Examples of these risks ranged from "Inflexibility in adapting the methodology to evolving project needs." which was score a 1 in likelihood and 3 in impact, to other risks such as "Lack of student engagement and participation" which was given a likelihood score of 3 and an impact rating of 4 so overall 12. This rating is very high and meant we kept more of an eye out for these risks and tried to avoid and minimise the risk they would cause, if they did arise. More examples can be seen in Figure(3):

Risk	**Likelihood (L): 1 to 5**	**Impact (I): 1 to 5**	**Total (L x I)**	**What can you do to mitigate the risk?**
Lack of familiarity with LEAN and Generative AI.	**2**	**4**	**8**	Conduct training and workshops for project team members to acquire the necessary skills and knowledge in LEAN and Generative AI.
Lack of student engagement and participation.	**3**	**4**	**12**	Implement strategies to motivate and engage students in the study, such as incentives. Schedule sessions at convenient times for students.
Lack of engagement when interviewing lecturers.	**2**	**3**	**6**	Schedule interviews at convenient times around lectures.
Technical challenges with computer systems.	**1**	**4**	**4**	Have a backup plan for alternative AI tools or methodologies if technical issues arise.
Bias in AI feedback.	**2**	**3**	**6**	Regularly monitor AI feedback for biases and take corrective actions. Ensure to use an AI with diverse data to reduce bias.
Limited University content.	**1**	**3**	**3**	Regularly consult with the maths faculty to align the project with the university's mathematical proof content to ensure relevance

Figure 3: Example of our risk analysis

3.1.2 Working as a group

The MTH3035 Mathematics Group Project is a valuable component of our curriculum, designed to cultivate essential transferable skills with enduring real-world applications [Uni23]. Recognising the importance of skills such as teamwork, critical thinking, leadership, project management, and effective communication, our group has approached this project with a keen awareness of improving and utilising such skills. Throughout our literature review we became aware of the profound benefits working in properly structure group can have on our project such as “Refining understanding through discussion and explanation”, “Tackling more complex problems than they could on their own and sharing diverse perspectives” [Uni23].

To address potential challenges associated with group work, particularly the issues of miscommunication, an uneven workload, time-management and decision making, we proactively ensured fair task delegation, considering each member’s strengths and weaknesses.

A fundamental element of our success as a group has been our commitment to ensuring effective communication [Cer14]. In our initial meeting, we collectively discussed and decided on various modes of communication that suited the diverse preferences of our team. An informal Whatsapp group was established for quick organisational updates, while we requested our supervisor to create a Microsoft Teams platform. This served as a centralised hub for communication, allowing us to upload and access project-related documents, such as Lean exercises and MTH1001 lecture notes. In order to maximise time with our supervisor we would discuss on the Whatsapp group weekly goals for individuals and specific topics of conversion needed to have in our weekly meetings, ensuring this list was in order of priority. Further, recognising the challenges of virtual communication, we opted for in-person meetings, utilising a free and accessible space that facilitated effective collaboration through laptop projection. We believed in-person meetings were optimal as we could get a personal opinion, challenge what is being said and get a deeper depth of understanding. We could all view and edit the same documents, as well as brainstorming on the whiteboard.

However, it is important to note, despite everyone’s commitment to attending all our scheduled meetings, there were some unforeseen circumstances such as illness or travel difficulties which resulted in some absences of members. Although, to mitigate any impact, we had a robust system of information sharing in place. When a member missed a meeting they were filled in via WhatsApp and also delegated the task for the following week so our timeline was not compromised. This practice continuously kept our entire team up to date and well-informed. Additionally, ongoing communication with our supervisor allowed us to utilise valuable expertise. Attending their office hour supported us in integrating Lean into our website. While communicating any important updates via email, we kept them updated on our ongoing work. This specific approach to communication and collaboration has been pivotal in navigating the complexities of our group project.

3.1.3 Timeline and Summative tasks

A well-structured timeline is a critical component of successful project management. It provides a visual road map, ensuring that tasks are distributed efficiently and deadlines are met. Our timeline for the MTH3035 Mathematics Group Project reflects a systematic approach to accomplishing various tasks and achieving summative assessments. This approach was benifical for us as a systematic approach helps improve efficiency and to stay on track with the tasks [RLH18].

The timeline and summative tasks are intricately linked, as the planned progression aligns with the deadlines for each summative assessment. This cohesive approach ensured that our group not only completed the required tasks but also maximised learning opportunities and demonstrated a holistic understanding of the project's objectives. Regular reviews and adjustments to the timeline facilitated adaptability to unforeseen challenges, contributing to the overall success of the project.

3.2 Our methods

3.2.1 Focus groups

The foundation of our methodology lies in the insights gained from preliminary focus groups with first-year mathematics students. Scheduled on Mondays, these sessions serve as a crucial starting point, providing an initial understanding of students' current understanding of mathematical proofs and their learning preferences. As the primary purpose of focus group research is to draw upon the perspectives, emotions, beliefs and experiences of participants in a way in which would not be feasible using other methods, for example observation, one-to-one interviewing, or questionnaire surveys [AG97].

These sessions investigated students' perceptions and feelings about Lean, offering a comprehensive view of their experiences and challenges. The use of pre and post assessments/surveys will facilitate the evaluation of our findings, enabling us to identify areas for improvement in students' understanding of mathematical proofs. By determining their baseline knowledge and incorporating their feedback, we aim to tailor our intervention tools effectively, ensuring that our subsequent initiatives align closely with the students' needs and preferences. Hence allowing us to collect both qualitative [GS13] and quantitative feedback [AL20] to utilise in the evaluation. Qualitative data including verbal conversations with the students and quantitative data such as the students rating their confidence in response to our questions.

3.2.2 Preliminary Survey

In addition to the focus groups, we emailed each first-year student a quick 5-minute survey to better understand their opinions and understand better what we should focus on, all surveys can be found in the appendix of this report, or here. Due to the anonymity of the survey, students were able to respond honestly without the fear of upsetting lecturers or facing repercussions for utilising AI technologies to complete their assignments [Dye17]. We posed a variety of questions to the students on their thoughts of Lean, the aspects they find most difficult, potential tools for us to use, etc. We also wanted to know what they thought about ChatGPT, when and if they had used it, and whether they would use it for Lean or proofs. The survey was highly effective and valuable as it gave us a plethora of data to evaluate and several new suggestions for how best to support students in developing their understanding of proofs and Lean.

Following the development of our tools, we conducted a second survey featuring similar questions, which can be found here or in the appendix. We allowed students sufficient time to actively engage with and assess the effectiveness of our tools, thereby gaining firsthand insights into their benefits. A detailed analysis of these subsequent reports is presented later in this document.

3.2.3 Preliminary survey with lecturers

In addition to conducting a survey with students perspectives, we recognise the importance of incorporating insights from experienced educators [HJC+02]. Consequently, we conducted a survey with the lecturers, asking them their view on how much their students struggle with Lean, if they believe additional tools were needed, if they believed our tools would be beneficial and also their views on AI tools as which can be found here or in the appendix. This survey wasn't as successful as the student survey as far less lecturers teach Lean and therefore less understand it and so felt they had inadequate knowledge to complete the survey.

After we had built our resources, a similar survey was sent to the lecturers to see what they thought and whether they believed the students would benefit from them. This Survey was used to help collect quantitative data for the evaluation on the tools.

3.2.4 Preliminary interviews with lecturers

In addition to our surveys, we conducted an interview with mathematics lecturer, Dr Marasingha, who has subject specific knowledge on proof and Lean. This was to gain a deeper understanding of their expectations regarding proof writing and the challenges they observe in students' comprehension. Interviews are one of the most common ways to collect qualitative data in research [S14]. Further, during our literature review the utility of interviews became apparent, as they help explain and understand the behaviour, experiences and opinions of the research topic. Also since the questions used are normally open ended the information collected will be in-depth and detailed [CMNL23]. These interviews served as a valuable source of guidance, helping us align our interventions with the academic standards and expectations set by the university. By tapping into these expertise, we sought to bridge the gap between theoretical knowledge and practical application, ultimately improving the educational experience for students. The information gathered from these interviews will be integral in shaping our approach to improving feedback mechanisms and enhancing students' understanding of mathematical proofs. Below is what we uncovered:

Dr Marasingha explains that it can be challenging for him to communicate and convey to students what thought processes he undergoes when he demonstrates proof writing. He later states that students of later years come to him asking for help with proofs and the problem normally is that they do not know basic definitions or terms. He says even though these definitions are taught in stage 1, the process of using these in proofs is something that does not always happen in stage 1.

He thinks one challenge of Lean is the syntax, for instance, putting a colon or a comma in the wrong place can trigger an often-unhelpful error message. Dr Marasingha believes this can be off-putting for students. He thinks another challenge for the students is that they believe their Lean errors are as a result of a syntax mistake, when in fact they often do not properly understand the mathematical proof they are attempting. He says it can be difficult to disentangle the two problems. He also believes students can 'absolutely hate' Lean when they first encounter it but begin to prefer it when they realise it does help them.

Dr Marasingha believes Lean helps students, mainly by highlighting to them the parts of proofs that they do not understand. It helps to show students what it is to structure a proof.

Dr Marasingha believes having an understanding of mathematical proof is to be able to both read and write a proof. Specifically, to write a comprehensive proof is to “have a reasoned argument starting from axioms or theorems using step by step logic to reach a conclusion”.

He says to be able to read proofs, you will be able to have the same thought processes of the proof writer. Additionally, if you are reading an incorrect mathematical proof, you will be able to pinpoint where exactly it has gone wrong and why it means the proof will collapse. Dr Marasingha says that to be able to write a proof, you are aware of how changing the conditions of a proof will affect the validity of the proof. For example, if you were to alter the hypothesis, you would see how this changes the structure and direction of the proof. He says understanding proofs on a deeper level means you can grasp the broad structure of proofs, so much so that you are able to talk through and explain your thought processes behind the proof.

Dr Marasingha says it is an extra challenge to incorporate the use of Lean into the curriculum. It can be difficult to ensure there is harmony between Lean and the lecture notes. He says Lean can be helpful to highlight the difficulties students have understanding proofs, for example difficulties in understanding the rules of logic. Dr Marasingha then describes his experience of learning proofs when he was a student, he explains that acquiring the proving skills were reliant on him absorbing information from the proofs demonstrated by the lecturer. He says having a more interactive learning style such as Lean allows students to become aware of the rules they do and don’t understand, by actively testing their proof writing abilities.

Dr Marasingha says students feel more motivated to learn Lean now than in previous years, particularly because students now are required to submit summative assessments related to Lean. A significant number of people attend his coding club sessions, reflecting students’ pro-activity when it comes to Lean. However, before this, students saw learning it as an extra task so were less likely to engage in it.

Dr Marasingha says there are ‘almost certainly’ holes in proof teaching. Dr Marasingha says that mathematics modules typically do a bad job of developing critical skills that are so common in other disciplines. For example, he believes there is not enough discussion of incorrect proofs in class.

He says students often ask for solutions because they are unsure if their answer is correct, but Dr Marasingha says by 2nd and 3rd year, students should be able to know for themselves if their answer is correct or not. He thinks this could potentially be resolved if there were more discussions including the criticism of proofs.

As a result of this discussion, we were able to gain a deeper understanding into the challenges lecturers face in teaching proofs and difficulties students face when learning proofs. To be specific, we gained a deeper understanding of the obstacles students must overcome when using Lean during the early periods of their undergraduate mathematics degree.

Using this information, we can apply these insights to the contents of our tools. For example, we can insert basic terms and definitions into our website to help resolve the problem of students not being aware of whether or not their proof is correct. Additionally, we have seen that a predominant issue is the syntax of Lean and its error messages being unhelpful, which encourages our resources to focus on error messages specifically. Given that we don't want to be curriculum specific with our tool, and the fact that Dr Marasingha mentioned the challenge of integrating Lean into the education curriculum, we will decide to create a more broadly focused tool. Finally, we intend to produce a interactive tool given the suggestion that more hands-on tools are beneficial, which we have also discovered in our literature review.

3.2.5 Creating our website, integrating Lean and AI

To evaluate what supports feedback for students learning proofs for the first time, we decided to build a website. The website is aimed to support a student's learning experience by providing examples and exercises within Lean, exam exercises, Lean error support, and AI prompts.

3.2.6 Website Structure

3.2.6.1 Format

These different sections are designed to address challenges identified through our research, where students often encounter difficulties. We aimed to present these sections in a user-friendly manner, to ensure the information is engaging and easy to understand. Tables have been used to clearly layout information. See Figure(4) below, a screenshot from the website of a table we have used to layout different propositional logic operations. The format has been positively received with one student saying the layout was "easy to understand" in a focus group.

Rule	Symbol	Lean Input	Connection	Meaning
And	$\wedge$	\and	$P \wedge Q$	($P \wedge Q$) is true if both P and Q are true otherwise false.
Implication	$\rightarrow$	\to, \r, \imp	$P \rightarrow Q$	If P happens then Q happens.
Or	$\vee$	\or	$P \vee Q$	($P \vee Q$) is true if either P or Q is true (or both) otherwise false.
If and Only If	$\leftrightarrow$	\iff, \lr	$P \leftrightarrow Q$	P happens if and only if Q happen.
True	$\top$	true	$p \rightarrow \top$	P is always true.
False	$\perp$	false	$p \rightarrow \perp$	P is always false.
Negation	$\neg$	\not, \neg	$\neg P$	$\neg P$ is the opposite of P. If P is true, $\neg P$ will be false and vise versa.$\neg P$ is shorthand for $p \rightarrow \perp$.

Figure 4: Propositional logic table

3.2.6.2 Examples and Exercises

In this section of the website, we have solely focused on propositional logic due to time constraints. In this section we have used examples and exercises taken from Mathematical Structures (MTH1001) problem sheets. Initially, we looked to incorporate Lean so that users could type the solutions directly into the website. Once realising this method was very complex, we decided to use a "Try in Lean" button which opens Lean in a separate tab, where the example is displayed. This method was a lot less involved while also gaining similar functionality. This saved us time, enabling us to present to website to students earlier than expected in our plan. Giving us both more time to evaluate and allow for any improvements to be made. Finally, we included an exam exercises page, specifically focusing on past exam questions. This is targeted to support student in their revision.

3.2.6.3 Lean Error Support

From our initial research, we found students struggled with understanding errors and how to remove them. To tackle this issue, we have introduced a table which provides the most common errors which a student would encounter, their meaning and steps to overcome the error. This section has been positively received, with multiple students saying this section was "helpful" within the focus groups.

3.2.6.4 AI Prompts

The AI Prompts page has been created to support students when using AI such as ChatGPT when learning proofs. We have specifically focused on propositional logic prompts once again due to time constraints. Providing these examples should allow students to use AI more effectively and understand how detailed the inputs should be to get successful outputs.

3.2.6.5 Building the Website

Once we had understood the structure of the website, we needed to understand how to build a one. Building a website consists of three coding languages: HTML; CSS; and JavaScript [MDN23d]. HTML is used for building the website's "structure," "and its content" [MDN23b]. CSS is used to format different "styles" [MDN23a]. JavaScript is used to build "interactivity to your website" [MDN23c]. We used JavaScript to build buttons which revealed solutions to answers when pressed. learning how to code in these languages was crucial in ensuring we could deliver a tool able to support student.

Once we successfully coded up a first draft of our website, containing all sections, we needed to put it online to then distribute out to students. To do so, we used GitHub, "a code hosting platform for version control and collaboration." (Github Docs) This Version control system allows you to make your website public, while also allowing you to control updates made.

3.2.7 Creating our additional tool

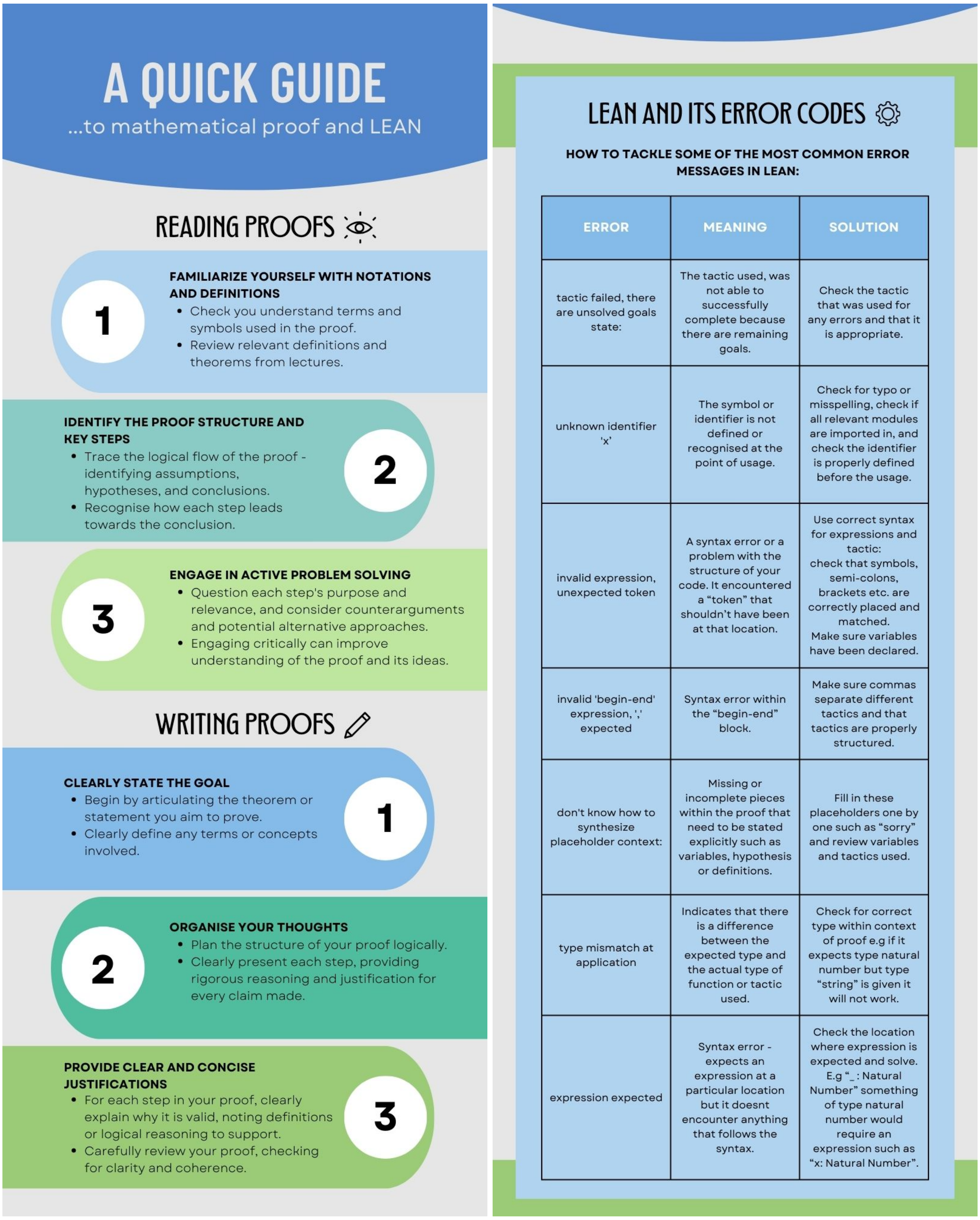
In addition to the creation of our website, we aimed to implement an extra resource to observe students' preferences of the two tools. After considering the potential different formats of resources, we decided on making an info-graphic which we would share to the students via the internet. Our team put together this info-graphic using Canva, an online graphic design platform. We included what we believed to be the most crucial information around proof writing and Lean, based on the evidence from our research into students' understanding in this area.

Following the student survey, we carried out some analysis in order to fully understand various aspects of our project. Dr Marasingha says it could be useful to create a resource that helps students understand error messages, made by hand instead of using AI. The resource could identify error messages and explain what they mean. Dr Marasingha also suggests using videos that work through proofs, illustrating best practices. He says it would be beneficial for a tool to connect Lean materials to lecture content and help show students how concepts learn in lectures are useful.us trends and patterns in the student's comprehension and to identify any knowledge gaps. Our findings were in line with expectations, and there was a strong consensus that students had difficulty writing proofs in Lean.

The first page of our tool focused on how to read a proof and also how to write a proof, which we discovered to be the two fundamental processes Dr Marasingha believes to be the foundations of understanding mathematical proof. These sections provided informative steps of the two processes - information we gathered in our previous research.

The second page of our tool presented a table of Lean error codes and explained where the error stemmed from and how to correct the problem. We integrated this table into our resource because it was clear from our research, such as the students' survey feedback, that a major struggle for students was understanding the unfamiliar syntax and error codes of Lean. We believed this table could assist the students when correcting their Lean code and hopefully cause Lean to be less off-putting for them.

We were then ready to present our info-graphic in the Coding Club and gather relevant feedback from students. Below, Figure(5) shows the leaflet we produced and presented to the students.



A QUICK GUIDE

...to mathematical proof and LEAN

READING PROOFS

1 **FAMILIARIZE YOURSELF WITH NOTATIONS AND DEFINITIONS**
- Check you understand terms and symbols used in the proof.
- Review relevant definitions and theorems from lectures.

2 **IDENTIFY THE PROOF STRUCTURE AND KEY STEPS**
- Trace the logical flow of the proof - identifying assumptions, hypotheses, and conclusions.
- Recognise how each step leads towards the conclusion.

3 **ENGAGE IN ACTIVE PROBLEM SOLVING**
- Question each step's purpose and relevance, and consider counterarguments and potential alternative approaches.
- Engaging critically can improve understanding of the proof and its ideas.

WRITING PROOFS

1 **CLEARLY STATE THE GOAL**
- Begin by articulating the theorem or statement you aim to prove.
- Clearly define any terms or concepts involved.

2 **ORGANISE YOUR THOUGHTS**
- Plan the structure of your proof logically.
- Clearly present each step, providing rigorous reasoning and justification for every claim made.

3 **PROVIDE CLEAR AND CONCISE JUSTIFICATIONS**
- For each step in your proof, clearly explain why it is valid, noting definitions or logical reasoning to support.
- Carefully review your proof, checking for clarity and coherence.

LEAN AND ITS ERROR CODES

HOW TO TACKLE SOME OF THE MOST COMMON ERROR MESSAGES IN LEAN:

ERROR	MEANING	SOLUTION
tactic failed, there are unsolved goals state:	The tactic used, was not able to successfully complete because there are remaining goals.	Check the tactic that was used for any errors and that it is appropriate.
unknown identifier 'x'	The symbol or identifier is not defined or recognised at the point of usage.	Check for typo or misspelling, check if all relevant modules are imported in, and check the identifier is properly defined before the usage.
invalid expression, unexpected token	A syntax error or a problem with the structure of your code. It encountered a "token" that shouldn't have been at that location.	Use correct syntax for expressions and tactic: check that symbols, semi-colons, brackets etc. are correctly placed and matched. Make sure variables have been declared.
invalid 'begin-end' expression, ',' expected	Syntax error within the "begin-end" block.	Make sure commas separate different tactics and that tactics are properly structured.
don't know how to synthesize placeholder context:	Missing or incomplete pieces within the proof that need to be stated explicitly such as variables, hypothesis or definitions.	Fill in these placeholders one by one such as "sorry" and review variables and tactics used.
type mismatch at application	Indicates that there is a difference between the expected type and the actual type of function or tactic used.	Check for correct type within context of proof e.g if it expects type natural number but type "string" is given it will not work.
expression expected	Syntax error - expects an expression at a particular location but it doesnt encounter anything that follows the syntax.	Check the location where expression is expected and solve. E.g "_ : Natural Number" something of type natural number would require an expression such as "x: Natural Number".

Figure 5: Our Handout - A quick guide to writing proofs

4 Analysis of findings

It was important for us to gain an initial understanding of how students felt about proofs and the use of Lean. Proofs are an essential topic within undergraduate maths and there is a deep emphasis on understanding the importance of proofs. it can be said that several students find proofs initially challenging, however, once students have had more practice and understand the proofs their understanding increases rapidly [SS08].

Figure(6) illustrates how diverse students' initial comfort levels with proofs are. With a mean score of 3, but a mode of 4, it appears that the greatest amount of students think they comprehend mathematical proofs moderately well, but with many also indicating a lower level of confidence. Looking at Figure(7), we see an even larger spread of perceived understanding, this time indicating a lack of familiarity with Lean. For this question, the modes are 1 and 3 which highlights the pupils' lack of understanding of Lean. This indicates that there is a considerable knowledge gap for Lean. This may even be hindering the students' understanding of proofs rather than enhancing it, as they seem to have a relatively high grasp of proofs compared to a much lower understanding of Lean. A similar trend presents itself whilst looking at Figure(8); understanding how beneficial students think Lean is in understanding proofs. This time the mode is 2 showing that students don't believe Lean will be a useful tool to develop their understanding of proofs. Previous research indicates that utilising software like Lean can be beneficial for students in the context of proof writing. It is reasonable to establish a connection between the shared characteristics observed in proofs created by students using Lean and their proficiency with the programming language [TI21]. We aim to change this and follow the trend of other studies and ideally, with the tools we develop and create, students will see how useful Lean can be and will opt to use it when writing proofs. Of all these graphs, Figure(9) displays the most pronounced opinion in responses, with 26 students stating that ChatGPT or AI is not effective in helping them understand mathematical proofs.

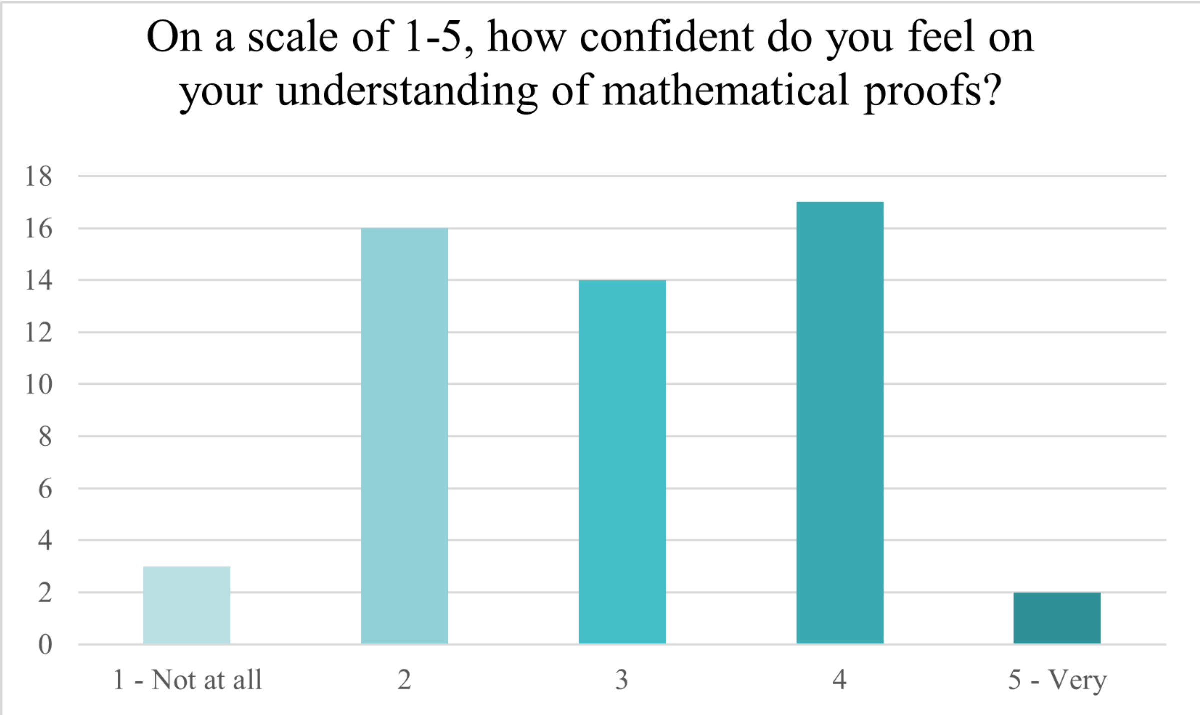


Figure 6: How confident students are with proofs

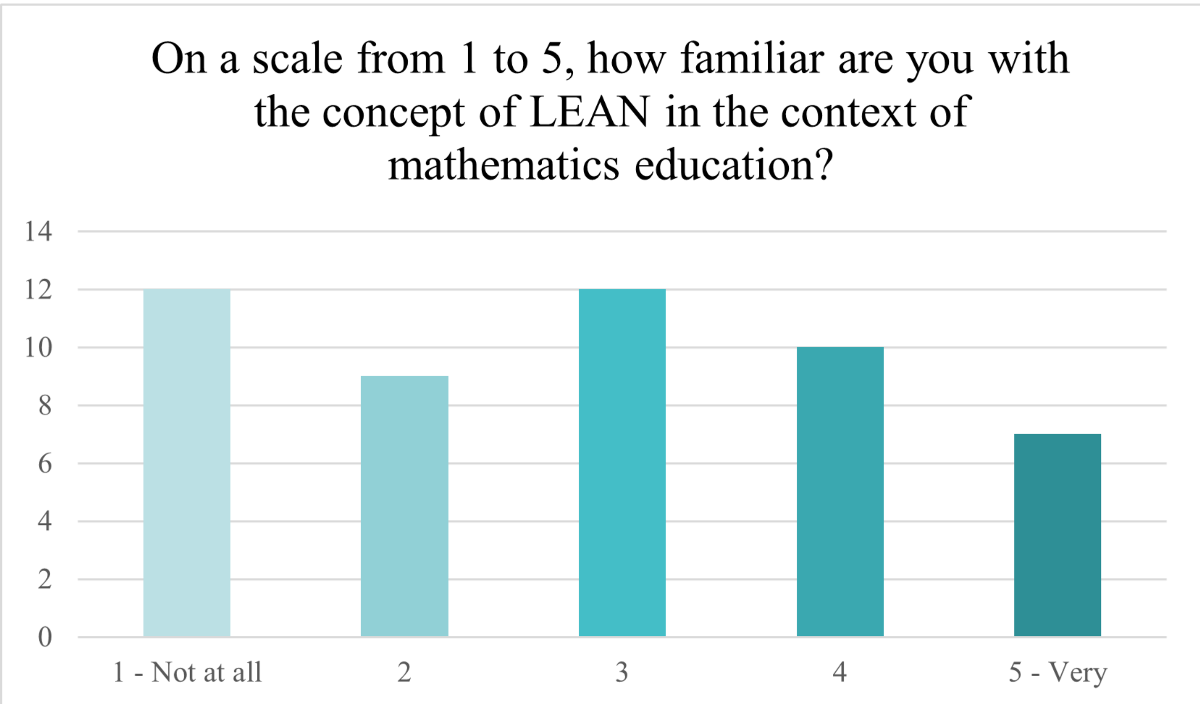


Figure 7: How familiar students are with Lean

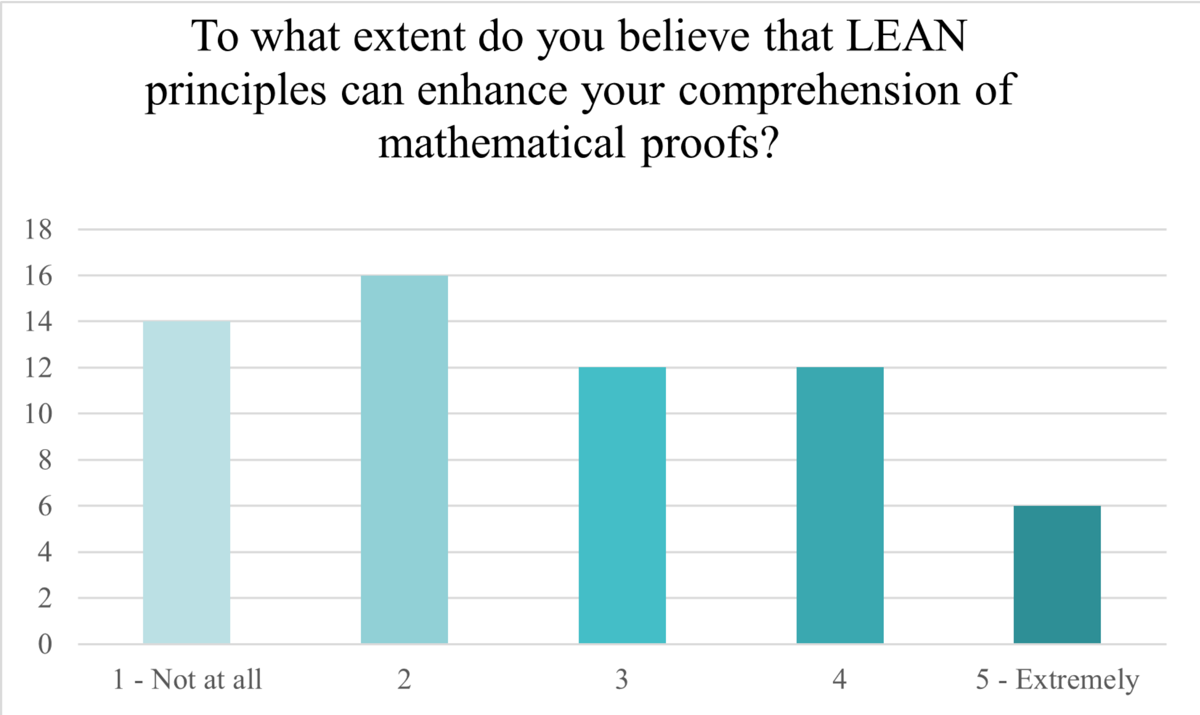


Figure 8: How beneficial students think Lean is to understanding proofs

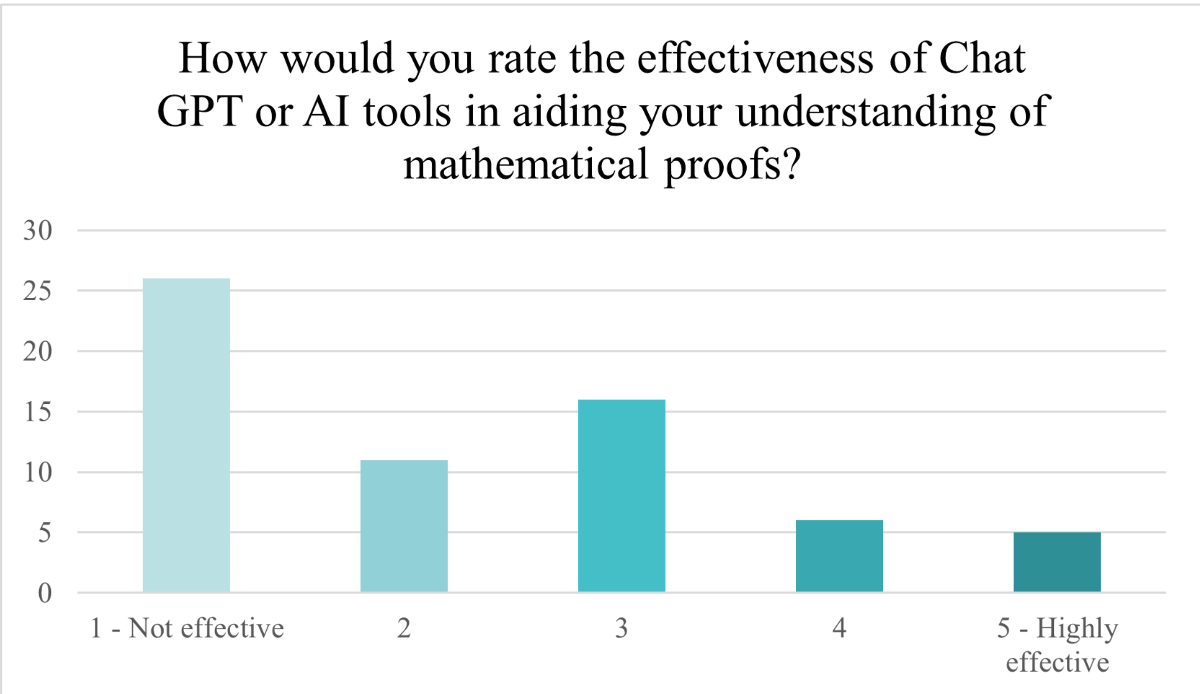


Figure 9: How effective students think ChatGPT is for mathematical proofs

Students were asked to respond to a survey regarding Lean in general, including whether or not they have used it, if it has benefited them and how. As can be seen in Figure(10), we coloured-coordinated these viewpoints according to their reaction. Positive views were indicated by green, negative by red, and neutral by blue. While some of the students' unhelpful opinions, like "I hate Lean" and "Used Lean to practice proving theorems," helped us understand their perspectives better, they did not provide us with a clearer picture of how to support them or where to concentrate. Other remarks like "I have, and the way proofs are formatted on Lean helped me come to a logical conclusion where I could understand HOW I proved the proof " This was more helpful because it made the problem easier to visualise and gave us a better understanding of what the students would benefit from when using Lean. Furthermore, because we now know the areas the students are having difficulty with, comments like "No, because it focuses too much on syntax" give us something to work on and an idea of how we can assist the students.

Overall, this open ended question provided us with a deeper understanding of specific opinions from students who have experienced using Lean alongside learning mathematical proof. The negative comments generally explained a disinterest and lack of understanding of the purpose of Lean within their education. Other helpful negative responses uncovered students' struggles to understand the specific syntax and error messages in Lean. The positive comments made apparent how useful the tool can be to improve understanding, in particular helping students structure proofs, and more easily visualise the end goals.

Figure 10: Students feedback of using Lean

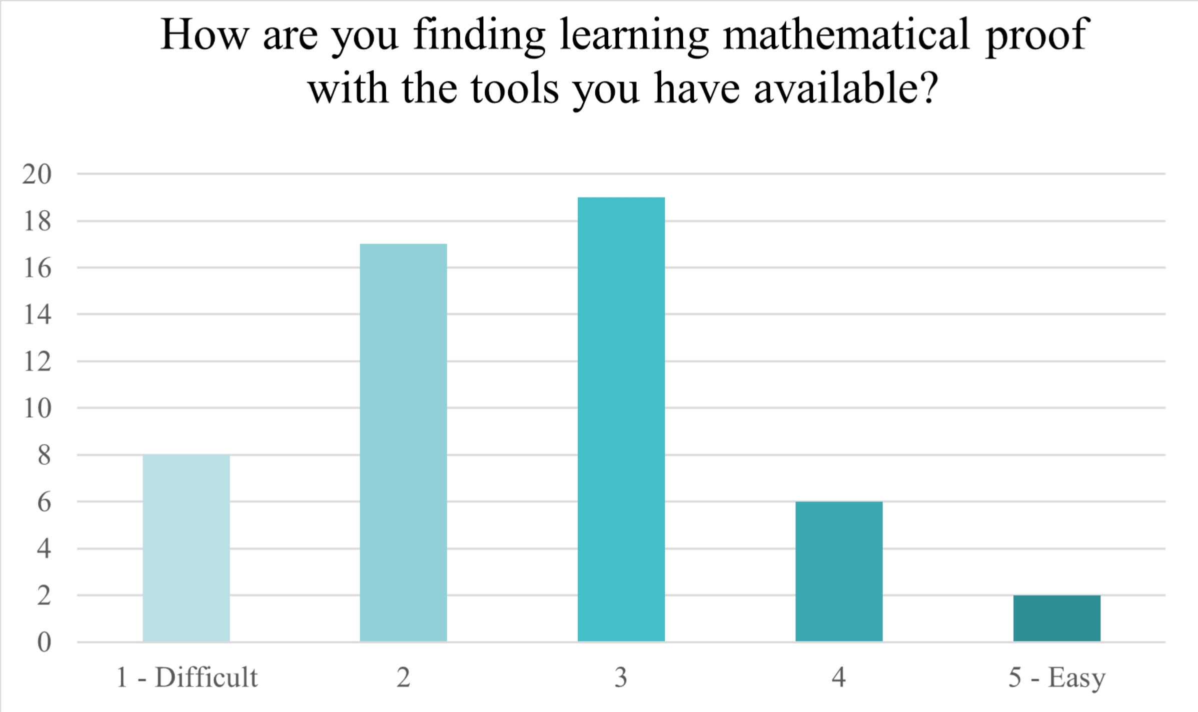


Figure 11: Effectiveness of any Lean tools already available to students

Moving on to the tools that students currently have at their disposal and their opinions about their effectiveness, we can see from Figure(11) that students believe their resources for mathematical proofs are inadequate and not aiding their learning enough. This makes our project even more important since it shows us the significant impact we can have and how urgently we should act. This is because the resources available to students reflect the students' performance, and the importance of these resources varies [Han06]. This motivates us to continue developing tools that will help students become more proficient in writing proofs and boost their confidence when using Lean.

After we had completed our tools, which we hoped would aid the students and alter some of the graphs results, we sent out a similar survey. This time, the students had access to both our website and our handout sheet and the survey asked students questions about their opinion on Lean with the aid of our tools and if they believe their skills with Lean have developed since using the additional tools. Our findings show that 75% of students believe that the website has positively impacted their ability to understand and work with mathematical proofs as seen in Figure(12).

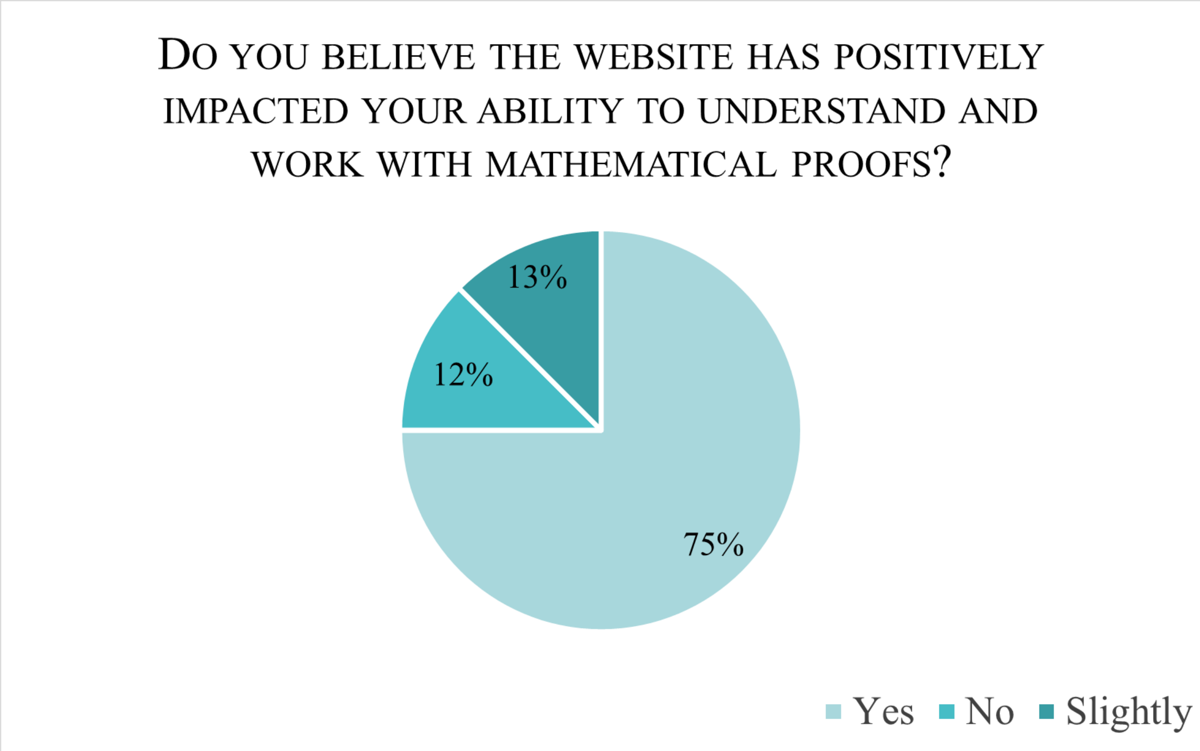


Figure 12: Pie chart showing the positive impact of our website

We can also see from Figure(13) that over 70% of students would recommend the website to a fellow student or colleague.

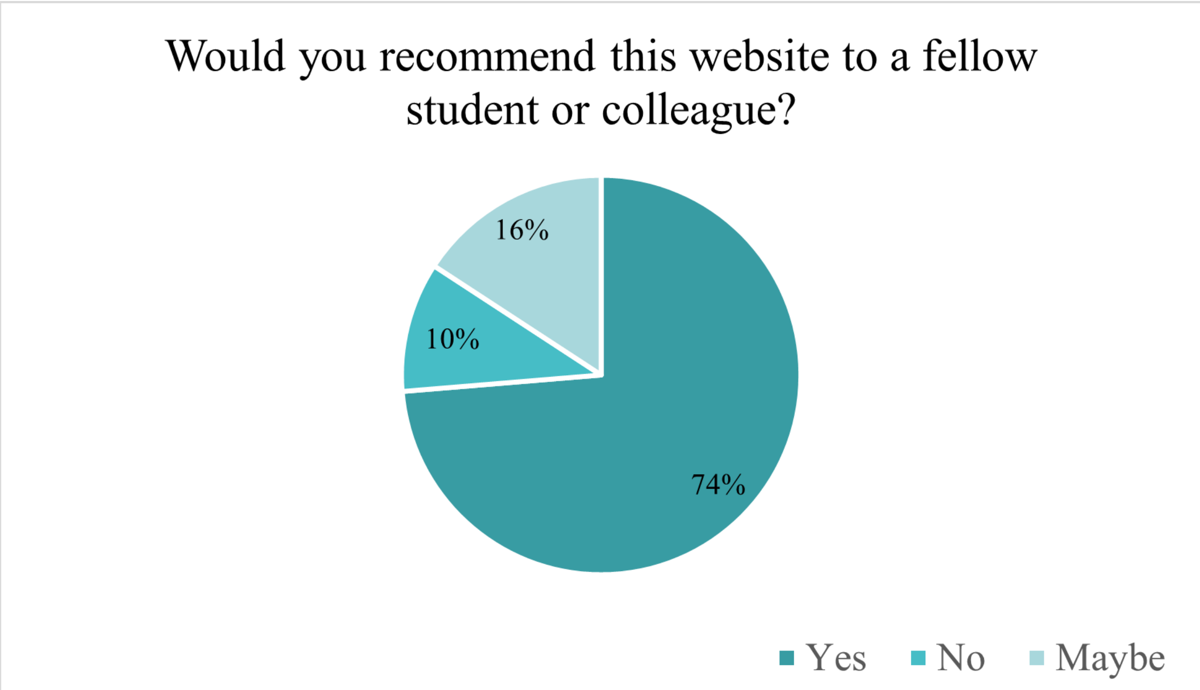


Figure 13: Pie chart showing recommendations of the website

Additionally, within the survey we asked the students how they believe our website could be improved, the majority of students suggested that more examples and solutions would be helpful and benefit them further.

Moving onto the handout, we also asked students similar questions, 60% of students gave a score of 4 out of 5 or greater suggesting that the handout contributes to their understanding significantly, this can be seen in Figure(14).

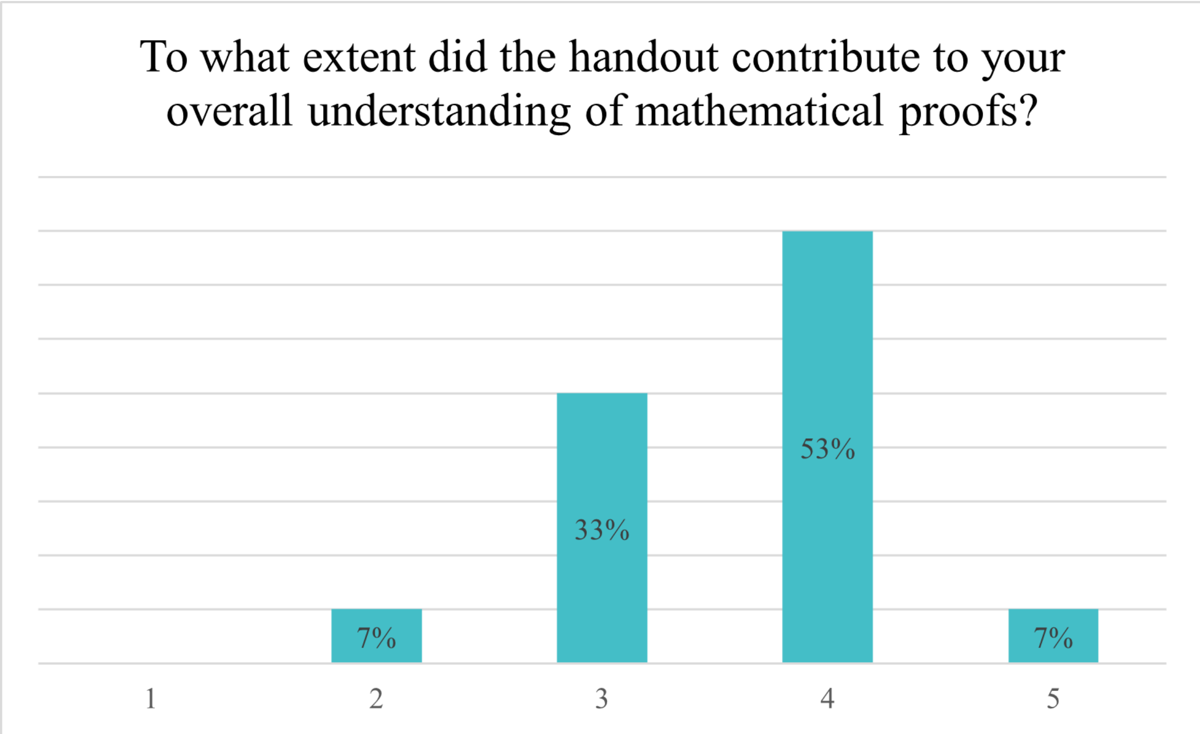


Figure 14: Chart showing improved understanding using the handout

We also asked the same question regarding the handout; would you recommend the handout to fellow students or colleagues. This time 86% agreed that they would indeed recommend the handout with the remaining 14% ‘maybe’ recommending it. This can be seen in Figure(15).

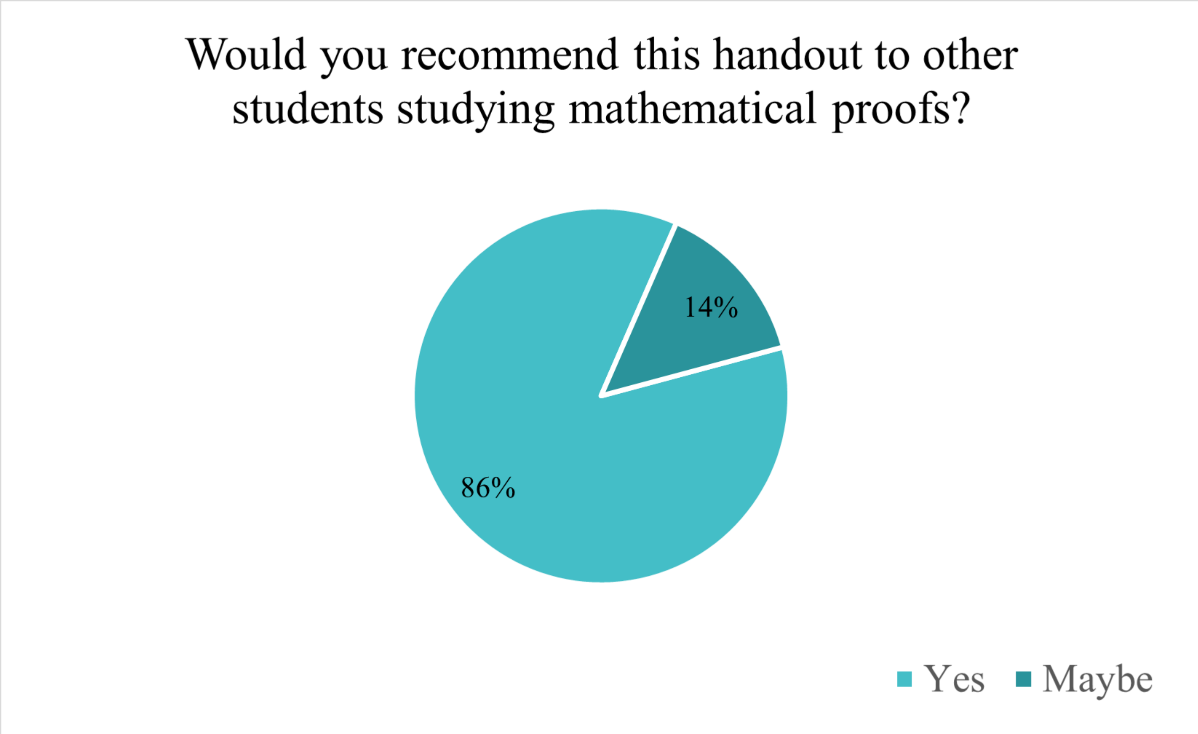


Figure 15: Pie chart showing recommendations of the handout

We asked the students in the survey what they thought could be done better with our handout. Students were free to offer feedback on any part of the handout, including the design and layout as well as the information it contained. The most frequently given response we got was to provide examples and a step-by-step guide for writing the proofs.

4.0.1 Analysing lecturers survey results

We conducted a comprehensive study of the lecturers' perspectives to ensure alignment with both student views and our own understanding, aiming to optimise our tools effectiveness. Their unique insights, stemming from teaching diverse student cohorts, some using Lean and others not, were invaluable. This approach was crucial for the pedagogical effectiveness of our tools, as we sought to integrate them into teaching practices with the active involvement and endorsement of lecturers, making their input vital to achieve our aims and objectives we laid out.

Even though a number of academics answered the survey, the number of replies was much lower than that of students. This is because the lecturers do not teach modules where Lean is required, therefore they are unable to remark on the strengths and weaknesses of their students.

We opted to analyse data by dividing it into two sections so that we can acknowledge any differences in the feedback or trends between the two sections, enabling a more comprehensive evaluation of the data. The first had the lecturer's feedback on students' use of Lean, and the second comprised the lecturers' thoughts on ChatGPT usage generally, but especially when it came to writing proofs and utilising Lean.

To start the analysis, we saw that 60% of lecturers thought that there wasn't enough support for students using Lean, seen in Figure(16). This was in concurrence with students beliefs and our own experiences as mathematics students. Hence demonstrating the exceptional need for intervention and the creation of new resources for the students struggling with Lean. This is supported by the conference paper 'Programming Assistance Software Tools to Support the Teaching of Introductory Programming'[KCC11] the paper highlights the challenges novice programmers face in learning and debugging, emphasising the importance to develop programming assistance tools designed to aid novice programmers. This would in turn increase students confidence in approaching Lean as well as ensure that lecturers feel comfortable that students have enough support in regards to using Lean[SFW14].

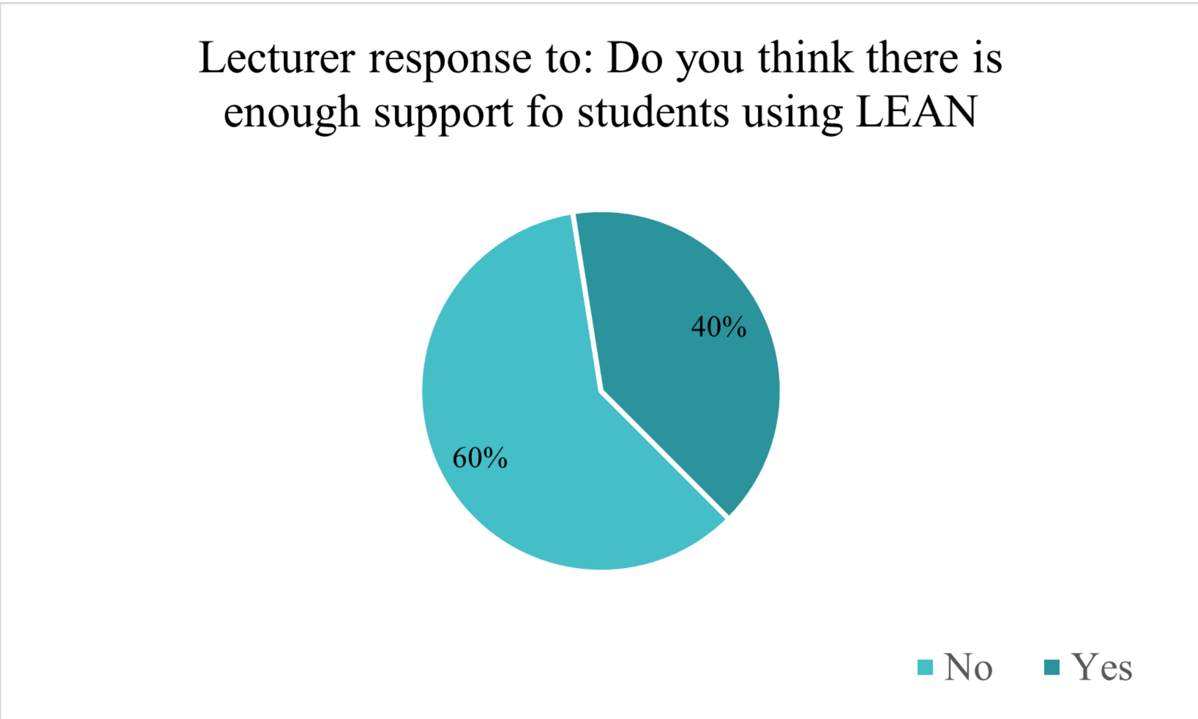


Figure 16: Lecturers view on support for students using Lean

Additionally, we were able to generate a word cloud that shows the lecturers' opinions about what they think will best assist students in becoming Lean proficient. We used a word cloud as in agreement the paper 'Exploring the Use of Tweets and Word Clouds as Strategies in Educational Research' from the University of Mauritius [CN22] the word cloud clearly visualises our data offering insights into the participants' perspectives. The larger the word, the more often it was proposed, as seen in Figure(17). This makes it easy to identify key themes, topics, or recurring terms at a glance.

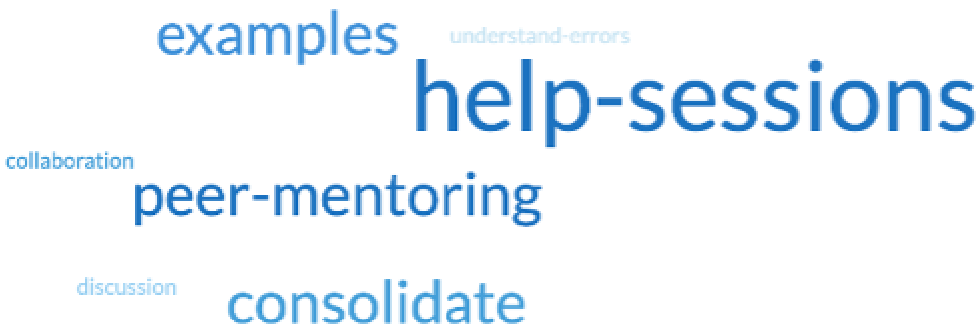


Figure 17: Word cloud, showing lecturers opinions on way to help students with Lean

We also obtained crucial data from lecturers, which strengthened our hypothesis that students would benefit from extra resources. First of all, lecturers believe that pupils had difficulty with Lean, rating it a 7-9 on a 10-point difficulty scale. However, it's important to note that while this fact holds true, all lecturers contacted unanimously stated that Lean can be an effective tool in helping students understand written proofs. Therefore, there is unanimous agreement among the lecturers that Lean is a valuable resource [TI21]. However, students often face challenges in understanding it, highlighting the essential need for additional tools. This is further in alignment with the findings of the article 'Interactive theorem provers for university mathematics: an exploratory study of students' perceptions', which, in their study involving 99 first-year students, concluded that Lean can be effective in bridging the gap to higher education but poses challenges in its syntax and perceived difficulty. [IT23]. As a group, we also concur with this analysis, being students ourselves who have learned Lean. We understand the initial challenges and daunting nature of the learning process, particularly an unfamiliar coding language. Consequently, we collectively agree that the development of more tools would prove immensely beneficial.

Additional analysis saw that 100% of lecturers agreed with our idea of creating a website with hints and prompts for Lean as well as a bank of keywords and their definitions would be useful for the students. Similarly 100% of the lecturers we surveyed agree that they think it would be useful for students to have a "How to guide" for using Lean. This consensus among lecturers is consistent with findings from Michelene T.H. Chi et al which also highlight the effectiveness of providing supplementary online resources for students learning Lean [MTC89].

By examining the second section of the survey, we can observe that lecturers oppose the use of ChatGPT for assignments and other assessed work. However, they are more receptive to the idea of utilising technology and AI when asked open-ended questions or when it comes to obtaining an idea of how to approach a problem rather than the actual solution. This finding was further consolidated by the conclusions reached by the University of Ahmad Dahlan on their study of “Exploring the Usage of ChatGPT in Higher Education” [Fir23]. However, in mathematics, it means that students lack understanding and miss crucial steps in learning [HQ98]. When looking in more detail about lecturers' opinion on ChatGPT for completing proofs the overriding idea was that ChatGPT does not understand proofs very well, and even less so than Lean [Fri23]. We also found this out by spending endless hours on ChatGPT understanding what result different prompts gave and how we could edit the prompts to ensure we were getting a refined and accurate result, however, this often wasn't the case. Lectures agreed it was better to use it in a methodical way and as a learning resource.

After creating our tools, we sent out a similar survey to lecturers to understand their perspective on the website and if they believed it would be beneficial. 100% of lecturers who answered the survey believed that the website would be beneficial and help students.

4.0.2 Analysis of Focus group with website

During the focus groups we introduced the students to our website and showed them where to find it if they didn't know about it yet. We put out Website on an ELE page that was available to all first year students studying MTH1001 Structures. We were able to learn about the student's opinions on the website and it's value after bringing our website to the focus group. We mostly questioned the students about their opinions on the website, asking them to note what they liked and felt would be helpful as well as what they didn't like and thought could be changed to make it even more helpful.

The general consensus was that it was very useful and would be beneficial with students saying “That's cool”, ”Oh that would be helpful”, “I think this could really help”. This was an expected and positive outcome, as the preliminary survey results show prior to the intervention of the website only 3.8% of students found learning mathematical proof easy with the tools they had available to them, whereas 84.7% of students rated finding learning mathematical proofs between a 1-3, where 1 is difficult and 5 is easy seen in Figure 11. Hence the introduction of a tool which would make learning mathematical proofs easier was expected to show a positives shift in these results. This was further consolidated by lecturers as seen in Figure 16 as 60% of lecturers who answered the survey believed there was not enough support for students using Lean, further demonstrating by introducing additional tools such as a website to offered further support to students it would be well received. The students especially like the practice questions with the answers so they could understand where they were going wrong and if they had made a silly mistake or if there was a fundamental issue with their understanding. Students having access to the answers to the practice questions we presented was essential for us as developing knowledge and understanding is based on monitoring oneself comprehension and misinterpretation. This type of learning leads to a better understanding of what they are learning [MTC89].

When asked how to improve the website, students made comments on the structure and format of the website, however they had fewer suggestions for improvement on the content. This is something that we agree given more time on the project, or the creation of an additional website, we would format it differently and create a more professional homepage. Learning from the Turkish Online Journal of Educational Technology's article on "User Satisfaction Evaluation of an Educational Website" we found that if we were to improve the usability and personalisation of the website we could greatly improve the number of website users.[Aki05]. Namely adding a log in function for the students to save their favourite resources and track their progress. Furthermore for example, we considered making the website into an app so our concept was more accessible to smart phones, as currently it is designed for laptops and desktop, leaving mobile users forced to turn the phone landscape due to the format.

From having the availability of the website along with the handout students opinions have gone from "I don't understand Lean" and "I can't understand the error codes" to "I can understand why Lean can be useful" and "I'm getting the hang of it more". This shows that the website has been a success and is a vital part of the students understanding and helps them to develop their skills further.

4.0.3 Analysis of Focus group with handout

We published our info-graphic on the Mathematical Structures ELE page to allow it to be accessible to students. We also presented it in the focus group and encouraged students to utilise it alongside their proof writing in Lean. This allowed us to collect crucial information on the practicality, usability and effectiveness of this tool. In this session we had students consolidate their apprehensiveness of the error codes provided by Lean, one said they 'don't know if the error messages mean anything' and the table on our handout makes these easier to understand. Many expressed they were in favour of the table because they found it translated the unfriendly error messages into a comprehensive text. We expected to gain this positive feedback on the table of error codes, specifically because we had already discovered that a huge hurdle of Lean is the misunderstanding of syntax and the error messages it provides. This is corroborated by Xiaoheng Yan and Gila Hanna's study "Using the Lean interactive theorem prover in undergraduate mathematics" [YH23] which addresses common questions that arise when utilising Lean. It explores its capabilities as well as its intricate relationship between syntax and semantics, although largely highlighting its effectiveness as a proof assistant.

A constructive point given by students was the emphasis of a more interactive resource as the online handout was not incredibly engaging. This was partially due to it being distributed to them online via the ELE page. Upon reflection, we would print off the handout and distribute it physically to the students as this has been proven to be a more effective learning method of learning from the "Screen vs. paper" study by the Karolinska Institutet University Library [MW15]. Furthermore, the leaflet did provide valuable material, although, it did not actively challenge the knowledge of students. Actively testing students' abilities is proven to be one of the most effective ways of teaching, not just mathematics, but a range of subjects [SFW14]. On top of this, students tend to prefer interactive teaching methods as opposed to more passive approaches, which was discovered in our research of teaching methods [CG22]. Therefore, we anticipated some negative feedback around the interactivity of our info-graphic and we agree in that a more interactive format, such as a website with an Interactive Theorem Prover would be more positively reviewed.

5 Evaluation

5.0.1 Evaluation of website

The evaluation of our website constituted a critical aspect of our research, aiming to assess its effectiveness in supporting students' comprehension of mathematical proofs. In the educational study, 'An assessment model for proof comprehension in undergraduate mathematics' published by Springer Link [MRFW+12b] the limited research on higher-level understanding and assessment of mathematical proofs at the university level is highlighted. The study emphasises that understanding a proof goes beyond the meaning and logical sequence of statements; it encompasses overarching ideas, key components, methodologies, and connections to specific examples. Consequently, the evaluation plays a crucial role in demonstrating the effectiveness of our website in improving students comprehension of mathematical proofs and provides valuable insights into the impact and usability of the educational tool developed during the project.

The website, designed to provide hints, prompts, and error definitions, underwent scrutiny. The survey results indicated positive shifts in students' opinions regarding Lean and skill development after utilising the website. This was expected due to the preliminary survey showing mass dissatisfaction with he current tools available and the desire for more interactive tools to meet the diverse needs of students. However, a critical analysis revealed areas for improvement. Some students expressed a need for more interactive elements, while others suggested additional clarity in certain sections. The user interface was generally well-received, but suggestions for enhanced navigability were noted, such as a tutorial on how to use the website so students used its full potential.

If our timeline extended we would have proceeded to of made those adaptations to our website. The effectiveness of the error message translation table was evident, demonstrated by the exit survey as 75% of students rated our Lean error code assistance feature on the website between as 7-10, 10 being the most useful.Yet the focus group highlighted that some students still found the syntax and language used in Lean challenging. This critical assessment illuminated the strengths of the website, such as its contribution to improved comprehension and error message understanding, while also informing iterative refinements to address identified areas for enhancement. Overall, the evaluation underlined the importance of an ongoing feedback loop to continually refine and optimise the website for maximum efficacy in aiding students in their journey to master mathematical proofs.

5.0.2 Evaluation of additional tool (Handout)

The evaluation of our additional tool, designed to improve students' understanding of mathematical proofs, involved a comprehensive and critical assessment. The importance of continuously creating and developing new tools is crucial in meeting the diverse needs of students alongside the evolving curriculum at university level education [Aki05]. Following our handouts implementation, we sought to evaluate its impact on students' proficiency with Lean and their overall grasp of proof-writing skills. The survey results showcased a positive trend, with students acknowledging an improvement in their skills since utilising the additional tool. However, a closer examination revealed significant insights.

One notable strength was the handouts ability to identify and translate complex error messages produced by Lean, a feature praised during the focus group discussions. Students expressed a increased comfort level with understanding and resolving errors, attributing this improvement directly to the tool's intervention. Nevertheless, challenges persisted, indicating that while the tool was beneficial, it did not entirely alleviate all difficulties associated with Lean.

Some students noted a preference for more personalised feedback, emphasising a desire for a dynamic tool capable of adapting to individual learning styles. This critical feedback highlights the importance of continual refinement to meet diverse student needs. Additionally, while the tool contributed positively to proof comprehension, some students expressed a need for further guidance in constructing rigorous proofs. This is something that we agree with and if we had the chance to create the tool again, this is something we endeavour to do.

The overarching evaluation of the additional tool revealed its positive impact on students' interaction with Lean and their ability to decode error messages. The critical insights gained from student feedback provide valuable direction for future enhancements, emphasising the iterative nature of tool development [Man18]. This evaluation reinforces the significance of creating adaptive and multifaceted tools that respond to students' evolving needs, thereby ensuring a robust and continually improving resource for mathematical proof education.

5.0.3 Evaluation of students mathematical proof comprehension with usage of additional tools

The evaluation of students' mathematical proof comprehension, with the usage of our additional tools, involved a multifaceted analysis that combined quantitative survey results, qualitative feedback, and a deeper exploration into the essence of understanding mathematical proofs. The survey findings demonstrated a positive trajectory, indicating an improvement in students' skills and confidence with Lean, as well as enhanced comprehension of proof-writing concepts. The findings were in line with our expectations as we strongly believed our tools would be of use and so were pleased when we saw this analysis.

The critical analysis of these results, however, revealed the need for ongoing refinement to ensure we were always developing something that students would find beneficial and effective. While the additional tools contributed significantly to decoding Lean error messages and fostering a more comfortable learning environment, some students expressed a desire for more personalised feedback and additional guidance in constructing complex proofs. This critical feedback shows the evolving nature of students' needs and emphasises the importance of continuous iteration in tool development. If we were to redo this project, we would create the same additional resources as we believe these are the most effective.

Qualitative feedback such as longer answer questions gathered during focus group sessions provided unique insights into the impact of the tools. Students appreciated the clarity brought to deciphering error messages and the overall improvement in their ability to approach proof-writing tasks. Yet, challenges persisted, highlighting the complexity of mathematical comprehension and the need for adaptive tools that cater to diverse learning styles.

To evaluate the magnitude of the impact on students' understanding of mathematical proofs, we revisited the fundamental question of what it truly means to comprehend a mathematical proof. Dr. Gihan Marasingha's definition, emphasising the construction of reasoned arguments through step-by-step logic, became the benchmark. While our tools had an overall positive impact on the students, challenges persist in translating this improvement across the spectrum of proof comprehension.

In conclusion, the evaluation of students' mathematical proof comprehension with the usage of our additional tools signifies a positive impact on developing students proof-writing skills. The critical analysis and qualitative feedback provide valuable insights for refinement, acknowledging the dynamic nature of students' development in understanding proofs. Our tools, while contributing substantially, highlights the ongoing challenge of achieving a comprehensive and universally improved understanding of mathematical proofs among students.

5.0.4 Evaluation of the interface between Lean and AI in helping students understand mathematical proofs

The evaluation of the interface between Lean and AI in aiding students' comprehension of mathematical proofs involved a focused analysis. Survey results demonstrated a positive reception, with students acknowledging the utility of the interface in guiding their understanding of proof-related concepts. However, critical analysis highlighted areas for improvement, particularly in tailoring AI feedback to address individual learning styles and providing more guidance in complex proof constructions.

The interface's effectiveness in decoding Lean error messages was a notable strength, contributing to an enhanced learning experience. Qualitative feedback demonstrated the importance of this feature, indicating increased confidence among students in navigating proof-related challenges [GS13]. It was important for us to use quantitative data as it provides a less bias opinion [Cra03].

Despite these positive aspects, the evaluation revealed that the interface's impact varied across different dimensions of proof comprehension. While students found support in decoding error messages, challenges persisted in constructing rigorous proofs. This nuanced insight calls for a more adaptive and comprehensive approach to the interface, acknowledging the diverse facets of mathematical proof comprehension.

In summary, the evaluation of the interface between Lean and AI demonstrated positive outcomes in specific areas of proof comprehension. The critical analysis and qualitative feedback guide future refinements [GS13], emphasising the need for a more holistic approach to enhance students' understanding of mathematical proofs.

6 Conclusions

6.0.1 Review of original aims and objectives

In conclusion, a comprehensive review of our original aims and objectives reveals significant progress in addressing the challenges students face in understanding mathematical proofs, particularly utilising Lean. Our research aimed to bridge the gap between the perceived difficulty of Lean and its potential as a valuable tool for comprehending written proofs. The unanimous agreement among lecturers on the utility of additional resources, including the website and handout, reflects a shared commitment to enhancing the learning experience for students. The integration of ChatGPT, while met with some reservations, underscores the need for thoughtful and methodical implementation of AI in mathematical education. As we evaluated the impact of our tools, positive shifts in students' attitudes and skills were evident by the evaluation of the qualitative and quantitative data we collected, emphasising the effectiveness of our interventions. However, the feedback also highlighted areas for improvement, reinforcing the evolving nature of teaching. Our dedication to continuous refinement, guided by both lecturer and student input, reinforces the importance of ongoing development in educational tools. Hence, this study reflects our accomplishments in this research and moves us forward on a path of continual innovation in facilitating a deeper understanding of mathematical proofs for students.

6.0.2 Future Directions of Research

Looking ahead, our research lays the foundation for several further avenues of research. First, recognising the strain on students' eyes during prolonged use, implementing a dark mode feature on the online website could contribute to a more comfortable learning environment, allowing for sustained concentration. We would research if this would have a positive impact on students when using the website.

Additionally, transforming the website into a mobile application would align with modern learning preferences, offering students greater accessibility and flexibility in engaging with the educational tools. This expansion to a mobile platform could potentially extend the reach and impact of our resources.

Furthermore, our handout, while effective, could benefit from an expansion to include a broader range of examples. This would cater to diverse learning styles and provide students with more varied illustrations, enhancing their grasp of mathematical proofs.

Beyond the immediate tools, exploring avenues to train AI in lean to overcome the current limitations of ChatGPT when asked to assist with Lean. By enhancing the capabilities of AI in understanding and guiding students through Lean-based proof construction, we can address current constraints and further improve the effectiveness of technological interventions in mathematical education. This would result in more accurate and instant feedback.

In conclusion, our commitment to ongoing research and innovation showcases our dedication to continuously refining and expanding educational tools. These future directions aim to adapt to evolving student needs, harnessing technology to create a more engaging, accessible, and effective learning experience in order to improve students comprehension of mathematical proofs.

References

[AAR23] Amos Azaria, Rina Azoulay, and Shulamit Reches. Chatgpt is a remarkable tool–for experts. arXiv preprint arXiv:2306.03102, 2023.

[AG97] Nigel Gilbert Anita Gibbs. What are focus groups? Social Research Update: Department of Sociology, University of Surrey, 1997.

[Aki05] Goknur Kaplan Akilli. User satisfaction evaluation of an educational website. Turkish Online Journal of Educational Technology, pages Article 11 p85–92, 2005.

[AL20] Robert Blizard Ingrid Tonni Chris Louca Albert Leung, Peter Fine. Teacher feedback and student learning: A quantitative study. European Journal of Dental Analysis, 2020.

[All97] Reg Allenby. Numbers and proofs. Elsevier, 1997.

[AP21] Thoma Athina and Iannone Paola. Learning about proof with the theorem prover lean: the abundant numbers task. International Journal of Research in Undergraduate Mathematics Education, pages 1–30, 2021.

[Arm82] J Scott Armstrong. The value of formal planning for strategic decisions: Review of empirical research. Strategic management journal, 3(3):197–211, 1982.

[Ave12] Terje Aven. Foundations of risk analysis. John Wiley and Sons, 2012.

[Bro16] Laura Broley. The place of computer programming in (undergraduate) mathematical practices. First conference of International Network for Didactic Research in University Mathematics, 2016.

[Cer14] H. Frank Cervone. Effective communication for project success. OCLC Systems Services: International digital library perspectives, 2014.

[CG22] Kojo Assuah Louis Osei Charles and Kwame Mantey Gershon. University mathematics major students' ratings of their lecturers' preferred teaching methods: A repeated-measures design. British Journal of Education,, 2022.

[CMNL23] Virginia Tech Carol M. Newman Library. Research methods guide: Interview research. University Libraries, 2023.

[CN22] Dorothy Cooshna-Naik. Exploring the use of tweets and word clouds as strategies in educational research. The Journal of Learning for Development - JL4D, 2022.

[Cra03] Duncan Cramer. Advanced quantitative data analysis. McGraw-Hill Education (UK), 2003.

[Des37] Rene Descartes. Discourse on Methode. 1637.

[Des41] Rene Descartes. Mediations On First Philosphy. 1641.

[Dye17] Harry T Dyer. Online anonymity: The good, the bad, and the ugly. HarryTDyer. com, 2017.

[EUR+17] Gabriel Ebner, Sebastian Ullrich, Jared Roesch, Jeremy Avigad, and Leonardo de Moura. A metaprogramming framework for formal verification. Proceedings of the ACM on Programming Languages, 1(ICFP):1–29, 2017.

[Fir23] Radha Firaina. Exploring the usage of chatgpt in higher education: Frequency and impact on productivity. Buletin Edukasi Indonesia, 2023.

[FK18] Sanna Erika Forsström and Odd Tore Kaufmann. A literature review exploring the use of programming in mathematics education. International Journal of Learning, Teaching and Educational Research, 17(12):18–32, 2018.

[Fri23] Pinchetti L. Griffiths R. R. Salvatori T. Lukasiewicz T. Petersen Frieder, S. Mathematical capabilities of chatgpt. arXiv preprint arXiv:2301.13867, 2023.

[GS13] Leonid Grebennikov and Mahsood Shah. Student voice: using qualitative feedback from students to enhance their university experience. Teaching in Higher Education, 18(6):606–618, 2013.

[Han06] Eric A Hanushek. School resources. Handbook of the Economics of Education, 2:865–908, 2006.

[HGS06] Fernando Hitt Harel Guershon and Annie Selden. Research in Collegiate Mathematics Education VI. American Mathematical Society, 2006.

[HJC+02] Richard B Hays, BC Jolly, LJM Caldon, P McCrorie, PA McAvoy, IC McManus, and J-J Rethans. Is insight important? measuring capacity to change performance. Medical education, 36(10):965–971, 2002.

[HM22] Fauziah Hakim and Murtafiah Murtafiah. Undergraduate students' levels of understanding in solving mathematical proof problem: The use of pirie-kieren theory. In AIP Conference Proceedings, volume 2575. AIP Publishing, 2022.

[HQ98] Larry Sowder Harel Quershon. Students' proof schemes: Results from explanatory studies. University of California San Diego, 1998.

[IT23] Paola Iannone and Athina Thoma. Interactive theorem provers for university mathematics: an exploratory study of students' perceptions. International Journal of Mathematical Education in Science and Technology, 0(0):1–23, 2023.

[KCC11] Melisa Koorsse, André Calitz, and Charmain Cilliers. Programming assistance software tools to support the teaching of introductory programming. 06 2011.

[KS21] Odd Tore Kaufmann and Børre Stenseth. Programming in mathematics education. International Journal of Mathematical Education in Science and Technology, 52(7):1029–1048, 2021.

[Kum05] Pankaja Pradeep Kumar. Effective use of Gantt chart for managing large scale projects. Cost engineering, 2005.

[Man18] Luke Mandouit. Using student feedback to improve teaching. Educational action research, 26(5):755–769, 2018.

[Man22] Manuscriptedit. Importance of staying up-to-date in research topics. Scholar Hangout, 2022.

[MDN23a] MDN. Css: Cascading style sheets. 2023.

[MDN23b] MDN. Html: Hypertext markup language. 2023.

[MDN23c] MDN. Javascript. 2023.

[MDN23d] MDN. Web technology for developers. 2023.

[Meg23] Chen Y. J Ferris J. A Knoth S Jones-Farmer Megahed, F. M. How generative ai models such as chatgpt can be (mis) used in spc practice, education, and research? an exploratory study. Quality Engineering, 2023.

[MRFW+12a] Juan Pablo Mejia-Ramos, Evan Fuller, Keith Weber, Kathryn Rhoads, and Aron Samkoff. An assessment model for proof comprehension in undergraduate mathematics. Educational Studies in Mathematics, 2012.

[MRFW+12b] Juan Pablo Mejia-Ramos, Evan Fuller, Keith Weber, Kathryn Rhoads, and Aron Samkoff. An assessment model for proof comprehension in undergraduate mathematics. Educational Studies in Mathematics, 2012.

[MRI09] Juan Pablo Mejía-Ramos and Matthew Inglis. What are the argumentative activities associated with proof? Research in Mathematics Education, 11(1):77–78, 2009.

[MTC89] Matthew W. Lewis Peter Reimann Robert Glaser Michelene T.H. Chi, Miriam Bassok. Self-explanations: How students study and use examples in learning to solve problems. Cognitive Science, 1989.

[MW15] Caroline Myrberg and Ninna Wiberg. Screen vs. paper: what is the difference for reading and learning? Insights: the UKSG journal, Jul 2015.

[Ort23] Sabrina Ortiz. What is chatgpt and why does it matter? here's what you need to know. ZDNET, 2023.

[oSC23] University of South Carolina. Importance of providing meaningful student feedback. Center for Teaching Excellence, 2023.

[RLH18] Peter H Rossi, Mark W Lipsey, and Gary T Henry. Evaluation: A systematic approach. Sage publications, 2018.

[S14] Jamshed S. Qualitative research method-interviewing and observation. Journal of basic and clinical pharmacy, 2014.

[SFW14] Miles McDonough Michelle K. Smith Nnadozie Okoroafor-Hannah Jordt Scott Freeman, Sarah L. Eddy and Mary Pat Wenderoth. Active learning increases student performance in science, engineering, and mathematics. Proceedings of the National Academy of Sciences Volume 111, Issue 23, 2014.

[SH16] Mamdouh M Soliman and Ahmed J Hilal. Investigating the effects of computer-assisted instruction on achievement and attitudes towards mathematics among seventh-grade students in kuwait. The International Journal for Technology in Mathematics Education, 23(4):145, 2016.

[SS08] Annie Selden and John Selden. Overcoming students' difficulties in learning to understand and construct proofs. Making the connection: Research and teaching in undergraduate mathematics, pages 95–110, 2008.

[Ste10] George A. Steiner. Strategic Planning. Simon and Schuster., 2010.

[TI21] Athina Thoma and Paola Iannone. Learning about proof with the theorem prover lean: the abundant numbers task. International Journal of Research in Undergraduate Mathematics Education, pages 1–30, 2021.

[TRI22] TRIARE. How to build an educational platform. Medium, 2022.

[Uni23] Carnegie Mellon University. What are the benefits of group work? Eberly Center: Teaching Excellence and Educational Innovation, 2023.

[Web08] Keith Weber. How mathematicians determine if an argument is a valid proof. Journal for research in mathematics education, 39(4):431–459, 2008.

[YH23] Xiaoheng Yan and Gila Hanna. Using the lean interactive theorem prover in undergraduate mathematics. International Journal of Mathematical Education in Science and Technology, 0(0):1–15, 2023.

Appendix

6.1 Links to surveys used in this research project

Preliminary survey for students: https://forms.gle/qcfiNBPexLaYL1EF8
Survey for students after usage of Tools: https://forms.gle/u5FrASNxqiMPLURF7
Preliminary survey for Lecturers: https://forms.gle/wa4BGmnu7UUMWj949
Survey for lecturers after usage of Tools: https://forms.gle/fMntktR2gcYJscD98

6.2 Link to the website

https://jwc220.github.io/Lean-Hub/

6.3 Code for the website we created

https://github.com/jwc220/Lean-Hub