



HAL
open science

Students' perceptions of a Proof Assistant, in an introduction to proof course, in the first year of University mathematics

Simon Modeste

► **To cite this version:**

Simon Modeste. Students' perceptions of a Proof Assistant, in an introduction to proof course, in the first year of University mathematics. Fifth conference of the International Network for Didactic Research in University Mathematics, Centre de Recerca Matemàtica [CRM], Jun 2024, Barcelona, Spain. hal-04944241

HAL Id: hal-04944241

<https://hal.science/hal-04944241v1>

Submitted on 13 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Students' perceptions of a Proof Assistant, in an *introduction to proof* course, in the first year of University mathematics

Simon Modeste¹

¹IMAG, University of Montpellier, CNRS, France, simon.modeste@umontpellier.fr

We have introduced a proof assistant, Edukera, at University in a mathematics course aiming at teaching proof. We investigate the perceptions of the students of this proof assistant, based on data gathered through questionnaires. We are interested in the effects of using a proof assistant on learning proof, but also on the utility, usability, and acceptability of such an interactive learning environment. Our results reinforce the idea that proof assistants can contribute to the learning of proof, and bring recommendations for the introduction of proof assistants in teaching settings.

Keywords: Teaching and learning of logic, reasoning and proof, Digital and other resources in university mathematics education, Computer assisted theorem proving, Students' perceptions, Edukera

INTRODUCTION

Proof is at the heart of university mathematics, and the teaching and learning of proof and proving is central issue at the beginning of university in mathematics (Hanna et de Villiers, 2012, in particular part V). In this context, the role of logic, and the interplay between syntax and semantics have been shown as important (Durand-Guerrier et al, 2012, Durand-Guerrier, 2008, Selden and Selden, 1995). Difficulties in the learning of proof has led to explore the potential of technologies in its teaching (Hanna et al, 2019), and the development of computer science raises issues regarding the interactions between mathematics and computer science (Durand-Guerrier et al., 2019). In this specific direction, the use of *Proof Assistants* for teaching has developed. Originally, proof assistants are expert software that help build and automatically check proofs, in mathematics and computer science. In a mathematics *introduction to proof* course in University, we have introduced and used a proof assistant called *Edukera* (<http://edukera.com/>), conceived for teaching proof and mathematics, and developed as an online exerciser. We generally aim at exploring and analysing the potential of proof assistants for the teaching and learning of proof and proving at University. More particularly, in this paper, we analyse students' perceptions of a proof assistant, through questionnaires given during two consecutive years. We will rely on the concepts of *Utility*, *Utilisability*, and *Acceptability* used in the study of interactive learning environment (ILE) (see Tricot et al., 2003). *Utility* concerns the learning efficiency of the ILE, *Usability* concerns the possibility to use it (manipulation, interface...). *Acceptability* concerns the representations regarding its utility and acceptability, conditioning the decision to use it. First, we will succinctly present proof assistants for education, and previous research. Then we will describe *Edukera*, and the teaching setting. Afterwards, we will specify our research questions, and methodological aspects. Finally, we will present and discuss our main results.

PROOF ASSISTANTS IN UNIVERSITY MATHEMATICS EDUCATION: CONTEXT, RESEARCH QUESTIONS, AND RESEARCH HYPOTHESES

Geuvers (2009) gives a general presentation and overview on proof assistants. Here, we rely on the description of a proof assistant (PA) given by Bartzia et al. (2022):

The term *proof assistant*, or *interactive theorem prover*, refers to a software tool allowing a user to interactively construct a formal mathematical proof. Some systems are designed to work in a specific domain such as geometry, logic or the analysis of computer programs, while others are general-purpose. Additionally, proof assistants used in the classroom can be sorted roughly in two categories: some are built by the community of educators and others are designed by specialists of interactive theorem proving for research or other professional purposes. (p. 254)

Many PAs have been developed, such as DEADuction, Lurch, Edukera for the first category, or Coq, Isabelle, Mizar, or LEAN for the second, and many of them are used or tested in educational purposes, in mathematics or computer science education. Although “we know almost nothing of [proof assistants’] potential contribution to other roles of proof, such as explanation, communication, discovery, and systematization, or how they now may become more relevant as pedagogical motivation for the learning of proof in the classroom” (Hanna et al., 2019, p. 9), it is increasingly admitted that using PA could contribute to the learning of proof.

Indeed, as tools that permit to build and check a formal proof, according to a logical system, PAs make explicit the formal rules governing the development of a proof, (such as the use of quantifiers and logic operators, and their manipulation in proofs). For instance, Chellougui (2020) shows that, out of a digital environment, introducing students to a logical system (Copi’s natural deduction in this case), “contributes positively to the students’ capacity to analyse mathematical proofs from the point of view of logical validity” (p. 319). But she identifies issues regarding the links that must be developed with the traditional forms of proofs in the mathematical course, and difficulties with the manipulation of the logical system for students. Using PAs could help overcoming these obstacles: as they offer an environment with (positive and negative) feedbacks in building a proof, and by taking in charge the checking of the validity of the steps of the proof, they permit to concentrate on the construction of the proof itself. As the tool focuses on aspects regarding the formal rules of the proof, they also allow to distinguish what concerns syntax from what concerns semantics, and could favour their dialectic in the development of proofs.

Although the idea of using PAs for teaching is not recent (e.g. Geuvers and Courtieu, 2007), there are still few experiments developed in a didactical research setting for the learning of proof in University mathematics. Thoma and Iannone (2022) compared the proofs produced by students engaged in a PA workshop (with LEAN) and other students. They observed two characteristics in the first group: accuracy of the use of mathematical language and proof writing resembling academic style, and division of proofs in goals and sub-goals. They hypothesize these characteristics to be

an effect of using the PA. In a different perspective, Bartzia et al. (2022) analysed five different PAs (including Edukera) for teaching, through the lens of one same classical exercise for the first year of university mathematics. Their a priori analysis investigated possible effects on student's learning of proof, and characteristics that are likely to strengthen or hinder these effects. For them, PA could possibly support:

- the reading, understanding and appropriate use of definitions;
- the reading, and control (and writing in some cases) of formal statements;
- the focus on the current proof state (possibly to the detriment of the structure and other parts of the proof) – with an exception for Edukera, where the entire text of the proof is always visible.

They mention the feedback not available in a paper proof, but stress a “tunnel effect”, amplified by automation, which may favour a focus on the current proof state, forgetting other parts, and allow a symbolic work to the detriment of sense-making.

In the line of these studies, we have experimented the introduction of a PA in a first-year course in mathematics. In this paper, we focus on student's perceptions of the PA, through the dimensions of utility, usability, and acceptability of the ILE. We address two research questions:

Q1 In which way do the students consider that using Edukera helped them in learning proof? and, on what aspects of proof do they find Edukera helpful?

Q2 Considered as an interactive learning environment (ILE) for the learning of proof in mathematics, how is Edukera perceived by its users, the students?

Based on the above discussions, we formulate the respective research hypotheses.

H1. Using an AP in university mathematics, as a tool that makes explicit (and controls) the logical rules governing the proof, can help students in:

- Understanding these logical rules, and their meaning (status of the statements, links with the logical operators and quantifiers...), and progressively distinguishing *truth* and *validity* (Durand-Guerrier, 2008);
- Identifying the syntactic and semantic dimensions of proof, being able to articulate them, and particularly connecting the syntactic forms of proofs to the structure of the statement to be proved;
- Understanding the global structure of proofs, and from this, developing skills in analysing, understanding, and writing proofs.

H2. First, the control by the PA on the validity of the proof steps, and the justification of all the goals, permits feedbacks (impossibility of some tactics, unsuccessful paths, unachieved remaining goals...) that help students in developing proofs and learning about them; and this automatic logical control allows any valid proof to be accepted by the PA, which can contribute to the learning of proof. Second, the exerciser format enables students to organize their homework as they want, and to progress gradually, possibly trying various solutions, which can foster their learning of proof.

We defend that these hypotheses strongly depend on links and alignment between the goals, contents, and proof styles in the course and the ILE.

EXPERIMENTAL CONTEXT AND THE EDUKERA PROOF ASSISTANT

Due to our objectives and some constraints, and in particular the need for an autonomous work from students on the PA, we have selected the Edukera platform. Edukera (Rognier and Duhamel, 2016) has been built for an educational purpose, and relies on Coq, a specialists' PA developed at INRIA (<https://coq.inria.fr/>). Edukera is no longer maintained but still available in its current state. It takes the form of an online exerciser where students have to build proofs of given statement, with a graphical presentation of the whole proof and a point-and-click interface (see figures 1 and 2). It has an imperative approach to proof, that is, “the user orders changes to be performed on the proof state (the current set of declared variables and constants, assumed hypotheses, and goals) using a predefined set of orders (also called tactics)” (Bartzia et al., 2022, p. 254). It proposes various pre-implemented proof-exercises, in witch teachers can make a selection, organized in a progression of notions (from boolean logics to set theory and functions) and difficulty. It includes tutorial exercises, where each new notion or tool is introduced. Edukera allows to choose various presentations of the proofs. We have chosen the mode called “Fitch”, which is the closest to traditional mathematics proofs (see figures 1 and 2). More technical details on Edukera an other PAs can be found in Bartzia et al. (2022). Figure 1 presents the user interface, and an exercise-proof for the theorem “the preimage of the intersection of two sets is the intersection of the preimages of the sets”. Edukera displays the variables declaration and the hypotheses on the top of the proof in construction, and at the bottom, the goal “to be justified”.

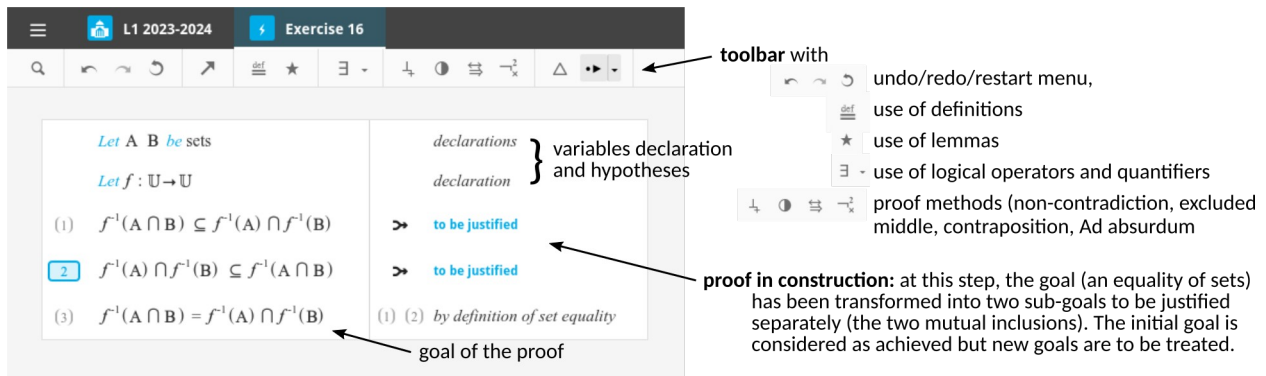


Figure 1: An exercise in Edukera, and the user interface.

As a first step, the user can work on the goal, using the definition of the set equality. Two new goals (the two mutual inclusions) appear as “to be justified” and the initial goal is considered as proved under those two new statements. The user can then decide to work on proving the second goal. Using the definition of the inclusion on this goal, and then the definition of the intersection on the hypothesis $x \in f^{-1}(A) \cap f^{-1}(B)$. Figure 2 shows what we get and possible following steps in the proof construction (right) in order to achieve it (no more statements to be justified).

This short example illustrates the way Edukera works, and the interactions a student can have with the proof in it. It will support our research hypotheses below.

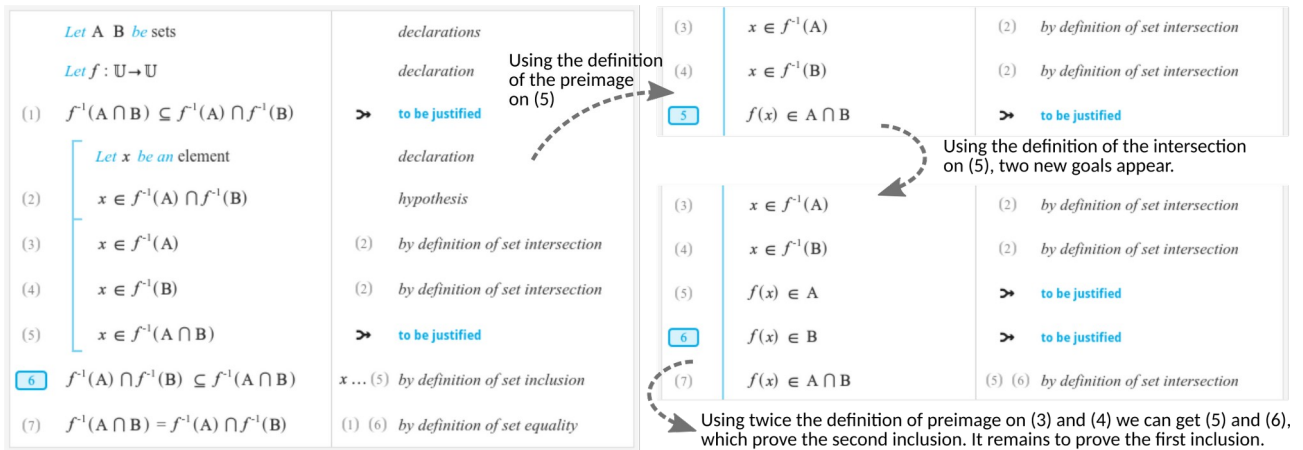


Figure 2: State of the proof in the following steps of construction (excerpts on right).

The use of the PA has been introduced in the first year of the bachelor of mathematics at the University of Montpellier (Kerjean et al. (2022) present various recent teaching experiences with PAs, including this one). The course, taught by the author of the paper, is a small course in the beginning of the year, called “Reasoning and set theory” which aims at introducing (before starting Algebra I and Analysis I) basics on the mathematical discourse, logics (boolean and first order logics), classical reasoning and proof techniques, notion of set (vocabulary and methods), and basics on functions (including injectivity, surjectivity, and bijectivity). The course attempts to make explicit the rules governing proofs, and aims to enable students to write proofs on elementary objects known since high school. This should enable them to better understand the presented proofs and to produce proofs in these courses.

The course is very small (9h lectures, 10,5h tutorial sessions) and concentrated on the three first weeks of the semester. Due to the limited teaching time, the PA has been incorporated as homework: it is introduced during the lectures in the beginning of the semester, and then the students are expected to work autonomously (the exerciser is structured to introduced features and concepts progressively, with tutorials). Two 1h extra lectures/demo/Q&A have been added in the semester. The PA has also been integrated as one of the three assessments, with the two first being traditional written tests. Students are expected to complete exercises regularly throughout the semester, and are evaluated on the number of exercises completed. We have made a selection of 170 exercises. Each proof-exercise can be attempted as many time as wanted, even when solved (while remaining considered as validated). This allows students to do (and redo) a lot of exercises. The grade obtained depends directly on the number of exercises validated (following a scale available in figure 5).

METHODOLOGY AND DATA COLLECTION

In order to study the perceptions of students according to these hypotheses, we designed a questionnaire to be completed by the students at the end of the semester, after the assessment. This questionnaire is composed of open questions, and Likert scale questions of agreement with some statements (Strongly agree / Somewhat agree / Somewhat disagree / Strongly disagree). The questions are presented in table 3.

Item	Questions	Type
1	Knowledge of Edukera and proof: I know how to use Edukera (for the type of exercises given this year) I know how to make proofs (for the type of exercises given in this course) I enjoyed using Edukera	Likert Likert Likert
2	Give one or more positive aspects of Edukera:	Open
3	Give one or more negative aspects of Edukera:	Open
4	On learning proof in mathematics: Using Edukera helped me better understand the structure of a mathematical proof Using Edukera helped me better understand how to write a proof in mathematics Using Edukera helped me learn how to start a proof in mathematics Using Edukera has helped me find ideas when I need to produce a mathematical proof Using Edukera has helped me better understand when I read a proof in mathematics	Likert Likert Likert Likert Likert
5	What Edukera helped me understand about proof in mathematics:	Open
6	What Edukera did not help me understand about proof in mathematics:	Open
7 and 8	Email if you agree to be contacted for research purpose, and free comment space	Open

Table 3: Content of the questionnaire

Item 1 permits to get auto-evaluation of students on proof, and the use of Edukera, and to know how much they appreciated working with the PA. Items 2 and 3 are designed to collect both positive and negative views on Edukera and its implementation in the course. Item 4 addresses the perception of benefits of Edukera regarding various dimensions of mathematical proof. Items 5 and 6 should give access more specifically to perceptions of difficulties and learning regarding proof. Item 8 can contribute in all those dimensions, by allowing extra comments.

This questionnaire was completed in December 2021 by 57 voluntary students (among 153) as a pilot study. Based on the preliminary results, it has been passed in 2022 in a mandatory format by 119 students (among 154). We will analyse here the answers from the two samples (n=176). We have also collected, for each year, the results of the whole students cohort (n=307), as a indicator of there investment in the PA and their achievements in it.

We will rely on quantitative and qualitative analyses, guided by our research questions and hypotheses. We will explore *Utility*, *Usability*, and *Acceptability* of the PA by using methods described by Tricot et al. (2003) as ILE's *evaluation by inspection* and *empirical evaluation*. Table 4 summarize the way we considered and evaluated the dimensions of utility, usability, and acceptability. Acceptability relates to the representations of utility and usability; in this sense, our empirical evaluation of students' perceptions informs particularly on acceptability.

RESEARCH RESULTS

Figures 5 and 6 summarize the principal quantitative data. In figure 5, we can see that more than 50% of the students solved 70 exercises or more in 2021 (mark 16/20 or more), and 60 exercises or more in 2022 (mark 15/20 or more). In 2021, only 17

students (11%), and 22 (15%) in 2022, did less than 25 exercises (marks less than 10/20), while more than 35% in 2021 (more than 25% in 2022) solved more than 90 exercises. We can notice that the two samples have quite similar profiles.

	Evaluation by inspection	Empirical evaluation
<p>Utility – learning efficiency of the ILE, matching between the ILE and the learning goals. <i>Focus on the learning of proof.</i></p>	<p>In this paper, this evaluation relates mostly to previous works, existing research (like Bartzia et al., 2022), and <i>a priori analyses</i> of PA, leading to our hypotheses regarding (Q1) for the Utility, regarding (Q2) for the Usability, and regarding both (Q1) and (Q2) for Acceptability.</p>	<p>Students’ perceptions of the learning benefits of the PA, mostly in items 4, 5, 6, and 8, but also 1 and students achievements.</p>
<p>Usability – possibility to use the ILE, its manipulation and its interface. <i>Focus on the PA itself.</i></p>		<p>Students’ perceptions of Edukera, as a digital tool, mostly in items 1, 2, 3, and 8.</p>
<p>Acceptability – representations about the ILE regarding its utility and acceptability, and conditioning the decision to use it. <i>Focus on the motivational aspects for using the PA.</i></p>		<p>Student’s perceptions of the learning benefits, and of the digital tool itself, and of the alignment between the course and the ILE. The motivational factors they mention. This can concern all the items. Achievements of the students (results and investment, see figure 5)</p>

Table 4: Dimensions of the PA’s evaluation, and contribution of the collected data.

It seems that the PA was rather usable: many students has involved in using it, and many exercises have been done. According to figure 6, in 2022, 80% of the students consider that they know how to use Edukera (73% in 2021), and 60% enjoyed using it (67% in 2021). This also contributes to the acceptability of the PA.

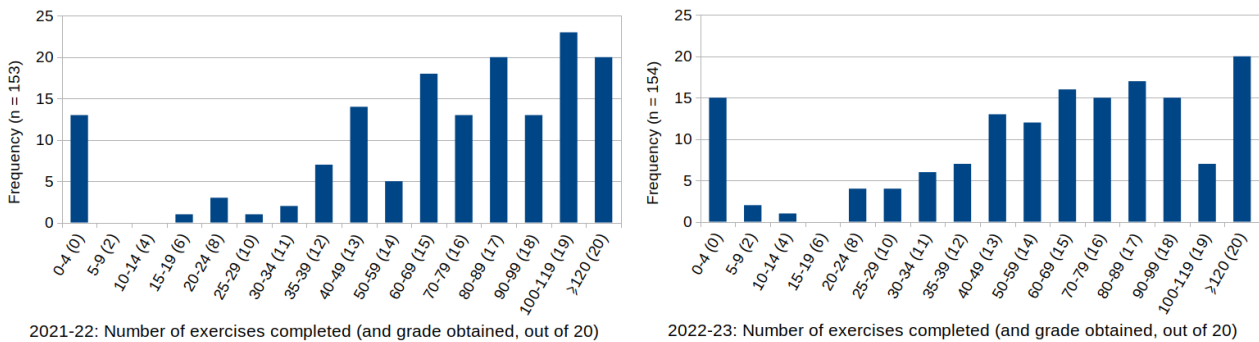


Figure 5: Distribution of the students by number of exercises treated (and marks).

Concerning Likert questions, we had quite similar distributions of the answers in 2021 and 2022, so we present here only the results for 2022 (higher number of answers). The answers to the diverse questions of item 4 are mostly positive, and rather few students strongly disagree. This supports the utility of the PA, and contributes to its acceptability. The two statements with the highest agreement are those related to understanding of the structure of the proof (70% agree, 29% strongly agree), and to the way to start a proof. Follows the statement about understanding when reading a proof. The two lowest (but still good) scores are for finding ideas, and

writing a proof. This hierarchy appears coherent with **H1**. Only the score concerning writing proof looks contradictory (44% in 2021 and 45% in 2022 disagree). We will discuss this later, in the light of the qualitative analysis.

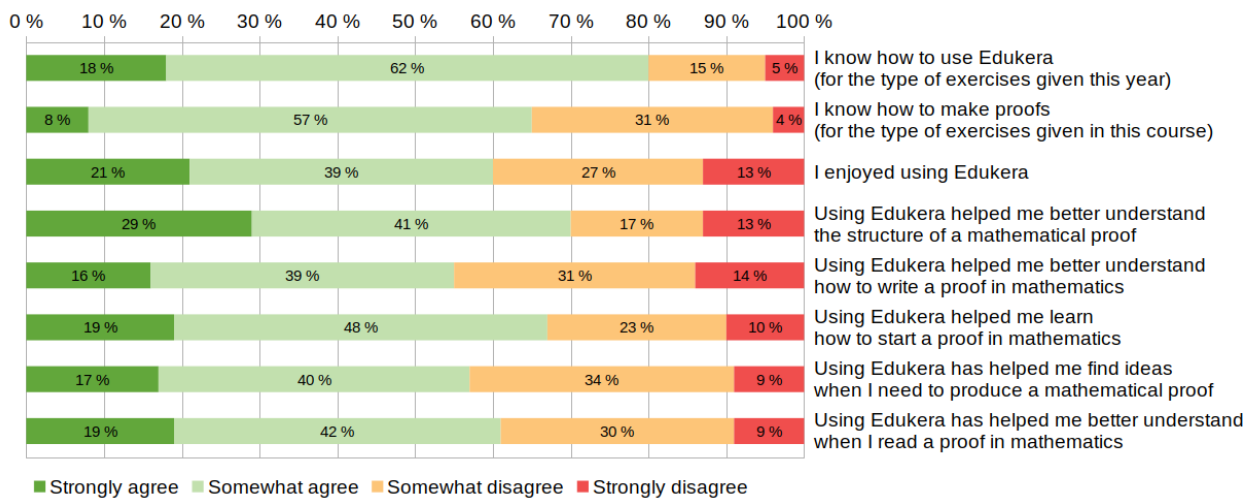


Figure 6: Students’ answers in Likert scales in 2022, items 1 and 4, on 119 answers.

About the qualitative analysis of the data (answers from 2021 and 2022), for this paper, we will limit ourselves to present and discuss our principal observations. Regarding utility, many students recognize benefits or interest of the PA for learning contents in link with proof (logics, methods, rigour, writing proof...). Some point out the interest of disregarding the writing issue to focus on the logical arguments of the proof. Among the negative aspects of Edukera, some students didn’t understand the mechanisms of the PA, or pretend that they knew how to make pen-and-pencil proofs of the exercises but did not manage to make the proofs with the PA, or consider that one can achieve many exercises by trial-and-errors without understanding. All these perceptions certainly contribute to the acceptability aspect.

When asked about what they learned from using the PA, students mention many aspects of proof: how to start a proof, the structuration of the proofs, how to write a proof, the meaning of operators and quantifiers, and the manipulation of hypotheses or other knowledge related to the deductive reasoning. Students even mention very precise notions, like scopes of the variables or specific type of proofs (inclusions, ad absurdum...). These results are coherent with **H1** and argue in favour of the utility of the PA. However, an element appeared more prominently than expected: many students say they learned about the role of the definitions in proofs, their importance, and the way they are used. A student express that he learned from the PA “the fact of going back to the definitions (and quoting them every time you use them, and that it is not trivial)”. This was especially present in the conclusions of Bartzia et al. (2022). Indeed, this is also about the logical status of the statements in the proofs, mentioned in our research hypotheses, and known by research as an key issue in learning proof.

Difficulties in learning with Edukera mainly concern linking pen-and-paper proofs with formal proofs, as already mentioned, and specific tactics, or proof methods like excluded middle, or ad absurdum, which are known as innately difficult for students.

About usability, some students found the PA simple to use once they got familiar with it. Some enjoyed the interface and its design, or the exerciser aspect (progressive and autonomous work, possibility to retry many times...). This relates to acceptability, and many motivational aspects are also mentioned: playful aspect, liberty and mobility in the work, quantity of exercises... The form of the assessment is also mentioned as motivating. Usability is also the dimension that concentrates most of the negative comments. The principal problems perceived by the students are: the difficulty to familiarize with the PA, the lack of explanations, and the many bugs in the software. This adds to a negative evaluation of the acceptability. Another acceptability issue mentioned by the students is the distance between the course and the PA, and in particular the vocabulary of tactics, and the representations of proofs. The need for alignment between the PA and the course was one of our points. This can explain the weaker score of the “writing proof” statement mentioned above.

Although the PA was only used in homework, the perceptions of students contains many positive points. Most of the criticism (bugs excepted) concerns the place of the PA in the teaching setting. It seems possible and reasonable to plan more tutorials to familiarize with the PA, to connect better the course with the PA and its features, and to consider work on links between formal proofs in the PA and pen-and-paper proofs.

CONCLUSION

Our results support the learning potential of proof assistants. The role of definitions has appeared stronger than expected in students’ perceptions, and must be further studied. One of our recommendations regarding the use of PA for learning proof is the need for a support to familiarize with the PA, and explicit links between the course and the PA. Our results are comparable to those of Iannone and Thoma (2023) on similar issues. They identified stronger difficulties than we did regarding the syntax of the PA (Lean). It can be due to the specificity of the PA, and research such as Bartzia et al. (2022) should be continued. This research is focused on students’ perceptions, which is not enough to validate the benefits of PA in learning proof. Further research is needed to address the effects on students’ proof and proving skills.

REFERENCES

- Bartzia, E., Meyer, A., & Narboux, J. (2022). Proof assistants for undergraduate mathematics and computer science education : Elements of a priori analysis. In M. Trigueros, Barquero, Berta, Hochmuth, Reinhard, & Peters, Jana (Eds.), *Proceedings of INDRUM 2022* (p. 253-262). <https://hal.science/hal-03648357>
- Chellougui, F. (2020). La déduction naturelle de Copi comme outil didactique pour l’analyse de preuves mathématiques. *Recherches En Didactique Des Mathématiques*, 40(3), 319-361.
- Durand-Guerrier, V. (2008). Truth versus validity in mathematical proof. *ZDM The International Journal on Mathematics Education*, 40/3, 373-384. <https://doi.org/10.1007/s11858-008-0098-8>

- Durand-Guerrier, V., Boero, P., Douek, N., Epp, S. S., & Tanguay, D. (2012). Examining the role of logic in teaching proof. In G. Hanna & M. Villiers (Eds.), *Proof and proving in mathematics education* (p. 369-389). Springer. https://doi.org/10.1007/978-94-007-2129-6_16
- Durand-Guerrier, V., Meyer, A., & Modeste, S. (2019). Didactical Issues at the Interface of Mathematics and Computer Science. In G. Hanna, D. A. Reid, & M. De Villiers (Eds.), *Proof Technology in Mathematics Research and Teaching* (14) (p. 115-138). Springer. https://doi.org/10.1007/978-3-030-28483-1_6
- Geuvers, H., & Courtieu, P. (2007). Proceedings of International Workshop on Proof Assistants and Types in Education (PATE07).
- Geuvers, H. (2009). Proof assistants : History, ideas and future. *Sadhana*, 34(1), 3-25. <https://doi.org/10.1007/s12046-009-0001-5>
- Hanna, G., & de Villiers, M. (2012). *ICMI Study 19: Proof and proving in mathematics education* (40). Springer. <https://doi.org/10.1007/978-94-007-2129-6>
- Hanna, G., Reid, D. A., & De Villiers, M. (Eds.). (2019). *Proof Technology in Mathematics Research and Teaching*. Springer. <https://doi.org/10.1007/978-3-030-28483-1>
- Iannone, P., & Thoma, A. (2023). Interactive theorem provers for university mathematics : An exploratory study of students' perceptions. *International Journal of Mathematical Education in Science and Technology*, 1-23. <https://doi.org/10.1080/0020739X.2023.2178981>
- Kerjean, M., Le Roux, F., Massot, P., Mayero, M., Mesnil, Z., Modeste, S., Narboux, J., & Rousselin, P. (2022). Utilisation des assistants de preuves pour l'enseignement en L1 : Retours d'expériences. *Gazette de la SMF*, 174.
- Rognier, B., & Duhamel, G. (2016). Présentation de la plateforme edukera. In J. Signoles (Ed.), *Vingt-septièmes Journées Francophones des Langages Applicatifs (JFLA 2016)*. <https://hal.science/hal-01333606>
- Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, 29(2), 123-151. <https://doi.org/10.1007/BF01274210>
- Thoma, A., & Iannone, P. (2022). Learning about Proof with the Theorem Prover LEAN: The Abundant Numbers Task. *International Journal of Research in Undergraduate Mathematics Education*, 8(1), 64-93. <https://doi.org/10.1007/s40753-021-00140-1>
- Tricot, A., Plégat-Soutjis, F., Camps, J.-F., Amiel, A., Lutz, G., & Morcillo, A. (2003). Utilité, utilisabilité, acceptabilité: interpréter les relations entre trois dimensions de l'évaluation des EIAH. In *Actes de la Conférence Environnements Informatiques pour l'Apprentissage Humain* (pp. 391-402), Strasbourg.