A TOOL FOR IDENTIFYING FUTURE PRIMARY SCHOOL TEACHERS' CONCEPTIONS OF PROOF

Mickael Da Ronch[1], Ismaïl Mili[2], and Cécile Ouvrier-Buffet[3]

[1]University of Teacher Education Valais
[2]University of Fribourg, [3]Paris-Est Creteil University, LDAR

This research report addresses an underexplored yet crucial issue in teacher education: prospective primary teachers' conceptions of mathematical proof. Based on a questionnaire administered to first-year teacher education students in French-speaking Switzerland, the study confirms the robustness and cost-effectiveness of this methodology and situates its findings within established research. The results highlight the functions future teachers attribute to proof and reveal discrepancies between their conceptions of mathematical rigor and the reasoning they mobilize when evaluating proofs, pointing to major challenges for teacher education.

CONTEXT AND RESEARCH QUESTIONS

Mathematical proof is widely recognized as a core component of mathematical activity and understanding (e.g. Hanna, 2000). International research documents persistent difficulties in learning and teaching proof at all educational levels and highlights a major rupture between secondary and tertiary mathematics education (e.g. Gueudet, 2008). Another critical transition, however, remains underexplored: the transition from secondary education to teacher education for prospective primary school teachers. Unlike secondary mathematics teachers, primary school teachers are generally not mathematics specialists and often experience difficulties in understanding and producing proofs (e.g. Barkai et al., 2002). They frequently lack confidence in their mathematical skills and struggle to manage classroom interactions and to foster a mathematical community during proving activities (e.g. Yackel & Cobb, 1996; Stylianides, 2016). Despite curricular injunctions and research showing that proof and argumentation can and should be meaningfully integrated into elementary mathematics instruction (e.g. Stylianides, 2016; Maher, Powell, & Uptegrove, 2010), research on future primary teachers' conceptions of proof remains limited. During teacher education, they simultaneously undergo an acculturation to mathematics and construct their professional identity as teachers. Understanding their conceptions of proof during this period is therefore essential for designing coherent teacher education interventions. Large-scale studies addressing conceptions of proof rely on resource-intensive instruments (e.g. Stylianou et al., 2015). Working at a local scale with fewer resources, this study examines whether conclusions from such studies are transferable to national teacher education contexts. It thus raises questions about the robustness of existing instruments: can a reduced but relevant subset of items yield coherent results?

2026. In H. Krzywacki, M. Hähkiöniemi, E. Haataja, M.S. Hannula & P. Hästö (Eds.). *Proceedings of the 49th Conference of the International Group for the Psychology of Mathematics Education: Research Reports*, pp. 1099-1106. PME.

Accordingly, this study develops a cost-effective questionnaire adapted from existing research to identify future primary teachers' conceptions of proof. We focus on their conceptions of proof, school experiences with proving, and perceived functions of proof.

THEORETICAL BACKGROUND

Our analysis draws on Balacheff's (1988) typology of arguments, particularly naïve empiricism, crucial experiment, generic arguments, and intellectual proof. We also select and adopt a functional perspective on proof, based on Hanna (2000): verification, explanation, systematization (formalization), discovery, and communication, with convincing added as a distinct function. Given documented student difficulties with deductive proof, the questionnaire targets fundamental components of deductive reasoning. It builds on and extends previous large-scale surveys by reusing validated items (Healy & Hoyles, 1998; Stylianou et al., 2015) and integrating new features as proposed by Ouvrier-Buffet (2023). In particular, it combines multiple-choice evaluation of written proofs with open-ended questions eliciting students' conceptions, including the types of proofs they know, the functions of proof they perceive, and their strategies to write, read, and evaluate proofs. This extension reflects the specific status of the population under study (freshmen, future primary teachers).

METHODOLOGY

Construction/design of the questionnaire instrument

The literature review by Ouvrier-Buffet (2023) identified only one study explicitly addressing proof conceptions among undergraduate students (Stylianou et al., 2015), building on frameworks developed by Healy and Hoyles (1998). These studies share Balacheff's (1988) framework to analyze proofs and rely on simple arithmetic and geometry problems. The present study extends this framework by integrating items on proof functions and evaluation strategies, reflecting the specific status of future teachers. The questionnaire relies on declarative data to ensure generality and to avoid dependence on advanced mathematical knowledge. It was designed to (i) minimize conceptual obstacles using familiar content, (ii) avoid unnecessary formalism, and (iii) reuse validated items. The questionnaire comprised four parts, administered in a one-hour paper-and-pencil format: Part I (global conceptions of mathematics), Part II (proof and multi-proof statements), Part III (school experiences and perceived functions of proof), and Part IV (demographic information).

Part II included two standard statements: (S1) "The product of three consecutive numbers is a multiple of 6" and (S2) "The sum of the angles of a triangle is a straight angle." For each, four written proofs reflecting different argument types and registers were provided, some incorrect or incomplete (see Table 1). The set of proofs was adapted from the large-scale study of Healy & Hoyles (1998). For S1, the proofs were adapted linguistically to the French-speaking context (ibid., p. 93, Kate's answer for A, Maria's answer for C, Nisha's answer for D) and enriched by the inclusion of a

crucial experiment in Answer B (ibid., p. 93, Leon's answer). For S2, the structure of Answers A, B and D was retained (ibid., p. 97, resp. Dylan's answer, Cynthia's answer and Amanda's answer), with adaptations aligned with francophone textbooks, and an additional visual proof of a generic nature that could be formalized was introduced for Answer C (see Figure 1).

Table 1: Features of the set of provided proofs (Part II)

	Arithmetic (S1)	**Geometry (S2)**
Ans. A	Natural language and arithmetic register. Correct proof.	Naïve empiricism / crucial experiment; numerical register.
Ans. B	Naïve empiricism and crucial experiment.	Intellectual proof. Formal language, alternating between natural and algebraic-symbolic registers. Correct proof.
Ans. C	Intellectual proof, formalized language in the algebraic register. Invalid reasoning. Incorrect proof.	Generic formalizable visual proof with natural language. Correct proof.
Ans. D	Intellectual proof, natural language, completeness of cases and disjunction of cases. Correct proof.	Naïve empiricism. Visual incomplete proof. Manipulation of cut-outs (common in francophone textbooks).

Figure 1: Excerpts of provided proofs (Part II)

S1, Answer B

$1 \times 2 \times 3 = 6$

$2 \times 3 \times 4 = 24 = 6 \times 4$

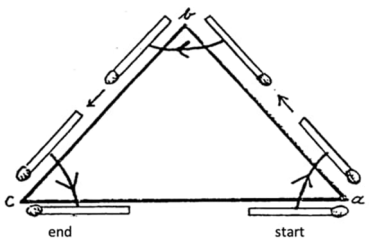
$4 \times 5 \times 6 = 120 = 6 \times 20$

$25 \times 26 \times 27 = 17550 = 6 \times 2925$

I can conclude that the product of three consecutive numbers is a multiple of 6.

S2, Answer C

Observe how the matchstick moves around the triangle. At the end, the matchstick is in the opposite position to the start. During this displacement, and by turning through the three angles of the triangle, the matchstick has made a half-turn.



(It comes from a Hungarian textbook)

Mixed methods of data processing and analysis

We adopted a mixed-methods approach combining qualitative and quantitative analyses. Drawing on assessment criteria adapted from Healy and Hoyles (1998) and Stylianou et al. (2015), we conducted a qualitative analysis of students' responses, complemented by quantitative coding, as detailed below.

Responses to open-ended questions in Part II, concerning known types of proof (Q1), functions attributed to proof (Q2), and criteria used to evaluate a proof produced by others (Q3), were analyzed using a combined inductive–deductive content analysis. Deductive categories were derived from the literature, while inductive categories emerged from the analysis of students' verbatim responses. Each of the 208 students' responses was coded according to the dominant category it expressed. The quantitative treatment then consisted in calculating frequencies and percentages in order to identify dominant categories within the population.

For Q2, coding focused on the functions attributed to mathematical proof, namely: explaining, communicating, verifying, convincing, formalizing, discovering, other, and no response/inability to identify functions. For Q3, the categories aimed to capture students' stances when evaluating a written proof. Five categories were distinguished: mathematical validity, comprehension/intelligibility, empirical checking, other, and no response/inability to evaluate. Responses classified as "other" referred to non-mathematical criteria used when reading a proof, such as quality of presentation, spelling, the source or authority of the author, or very general subjective judgements. These responses were distinguished from non-responses, as they expressed identifiable criteria that did not fall within the other categories.

The analysis of the multi-proof statements at the end of Part II was also category-based. It relied on Balacheff's (1988) levels of proof, the level of language used (formal versus natural), and the mathematical validity of the proof (correct/incorrect). Students' proof productions were first analyzed according to the identified level of proof. Subsequently, responses to the five questions related to Statements 1 and 2 were analyzed quantitatively to identify overall trends.

Finally, Part III of the questionnaire was analyzed quantitatively. This analysis made it possible to identify students' prior experiences with proof during their schooling, the importance they attribute to proof in the teaching and learning of mathematics, and the functions of proof they are able to identify in a controlled manner.

The questionnaire was administered at the beginning of the first year of a primary teacher education program (Bachelor's level) at two Swiss institutions: the University of Teacher Education Valais and the University of Fribourg. The sample consisted of 208 first-year students (86 from Valais and 122 from Fribourg), all entering initial teacher education after completing comparable secondary-level mathematics curricula.

RESULTS

Proof conceptions: known proof types (Q1), functions of proof (Q2) and control strategies (Q3)

Approximately 43% (n=89) of students reported not knowing different types of proof (Q1), while 17% (n=36) gave no answer and 2% (n=4) expressed uncertainty. Among respondents, many explicitly stated that they did not understand what a mathematical proof is. Several confusions emerged, with proof being assimilated to a known

theorem, a calculation, or conflated with everyday, legal, or historical meanings of "proof". Only a very small number of students mentioned canonical proof methods, such as proof by contradiction or recursion. Students' responses indicate a pluralistic but weakly stabilized view of proof (Q2). The functions of explanation and convincing (17% each, *n*=36), communicating and formalizing (16% each, *n*=34), and verifying (14%, *n*=30) were relatively evenly distributed. However, 14% (*n*=29) were unable to identify any function, 3% (*n*=7) referred to an "other" function, and discovery was almost absent (1%, *n*=2), suggesting that proof is rarely conceived as a means of producing new mathematical knowledge. Students typically identified only one function, pointing to fragmented conceptions shaped during prior schooling and weakly connected to proof's epistemological role. Regarding control strategies (Q3), the most frequent criterion concerned mathematical correctness: 39% (*n*=82) reported checking calculations ("if the calculations are correct and logical"). A smaller proportion referred to intelligibility or understanding (14%, *n*=29), while empirical checking through examples was marginal (2%, *n*=5). By contrast, 24% (*n*=49) relied on non-mathematical criteria (e.g. presentation, spelling, author credibility), and 21% (*n*=43) stated they did not know what to check. Overall, 44% (*n*=92) showed either no clear control strategy or reliance on non-mathematical criteria. These results suggest limited engagement with proof verification, likely reflecting scarce experience with authentic proof-related activities (searching, writing, checking a proof, mobilizing heuristics) during secondary schooling.

Proof conceptions: students' writings in multi-proof statements

Unlike the questionnaires of Healy and Hoyles (1998) and Stylianou et al. (2015), the present instrument included blank spaces allowing students to write arguments or comments. A majority did not write a proof for either the arithmetic statement S1 (52%, *n*=108) or the geometry statement S2 (61.5%, *n*=128). Among the 100 students who attempted a proof for S1, naïve empiricism predominated (*n*=58), typically through one (*n*=37), two (*n*=9), or three numerical examples (*n*=9). Only three responses approached a crucial experiment using larger values. Eleven students attempted syntactic arguments by translating natural language into formal expressions, but without engaging in semantic or deductive reasoning. More generally, when producing or evaluating proofs, students did not engage in logical–mathematical validation of arguments or structure, but relied on superficial plausibility judgments. "Verification" was understood as whether a statement appeared convincing rather than deductively justified. No interaction was observed between students' own productions and the provided proofs: neither were given proofs used to refine students' arguments, nor were personal attempts used to interrogate the proposed proofs. In geometry, most students who wrote a proof (79%, *n*=63 of *n*=80) translated the statement into a formalized language, sometimes with a drawing but without instruments or angle coding. The remaining students relied on naïve empiricism, summing numerical values of angles ("*60+60+60=180*" or "*23+47+110=180*") without checking or justifying them.

Evaluation of proofs for multi-proof statements

Overall results are summarized in Table 2. Most students identified proofs based on naïve empiricism or crucial experiments as closest to what they would have produced themselves: Answer B in arithmetic (60%, *n*=125) and Answers A (46%, *n*=96) and D (36%, *n*=74) in geometry. Similar patterns appeared for proofs chosen to explain the solution to a classmate (S1: Answer B, 47%, *n*=97; S2: Answer D, 51%, *n*=106). When asked to identify incomplete proofs, students selected similar responses in arithmetic (Answers A: 50%, *n*=104; and B: 57%, *n*=118), whereas in geometry they most frequently identified the visual, formalizable proof (Answer C: 56%, *n*=117) as incomplete. Intellectual proofs were predominantly identified as the most "thorough" or rigorous, even when they were mathematically incorrect (S1: Answer D, 54%, *n*=112; Answer C, 40%, *n*=84; S2: Answer B, 65%, *n*=136). Proofs judged as correct were mainly intellectual proofs (S1: Answers C, 73%, *n*=152, and D, 63%, *n*=130; S2: Answer B, 72%, *n*=149), although the naïve manipulative proof in geometry also received substantial endorsement (Answer D: 59%, *n*=123).

Table 2: Students' choices of proof (S1; S2) (*n*=208) (Part II)

Questions \ Answers	**A**	**B**	**C**	**D**
Which answer is nearest to what you would have done?	(42; **96**)	(**125**; 32)	(36; 12)	(6; **74**)
Which answer would you choose to explain the "solution" to your classmates?	(**58; 58**)	(**97**; 32)	(42; 33)	(17;**106**)
Which answer is the most "thorough" for you?	(8; 26)	(9; **136**)	(**84**; 18)	(**112**; 23)
What is/are the correct proof(s) for you?	(70; **94**)	(79; **149**)	(**152**; 70)	(**130**; **123**)
What is/are the incomplete proof(s) for you?	(**104**; **89**)	(**118**; 26)	(28; **117)**	(13; 57)

These findings reveal a clear gap between students' spontaneous conceptions of proof (largely aligned with naïve empiricism) and their recognition of intellectual proofs as mathematical standards of correctness and rigor. Notably, incomplete or naïve manipulative proofs (S2, D) were nevertheless considered both thorough (*n*=23) and correct (*n*=123), as was the visual, formalizable proof (S2, C), which, although unfamiliar in the Swiss context, was judged correct by a substantial number of students (*n*=70). This variability highlights unstable conceptions of proof across mathematical domains and a limited spontaneous mobilization of deductive reasoning.

Classroom experiences with proof and functions of proof

Part III show contrasting experiences. A large majority reported that proofs were often produced solely by the teacher (63%, *n*=130), while only a minority reported constructing proofs collaboratively with classmates (40%, *n*=84). Proof thus appears to have been experienced mainly as something to observe rather than to do. While 41%

agreed that proof was a central focus of mathematics lessons, and 78% (*n*=163) viewed proof construction as fundamental, opinions were divided on whether one can study mathematics without being skilled in proof. Overall, future primary teachers value proof highly but report limited opportunities for active engagement during schooling.

In a controlled format (*n*=208), students felt legitimate selecting multiple functions, unlike in the open-ended Q2. Verification (95%, n=198) and explanation (93%, *n*=194) were overwhelmingly recognized, followed by convincing (70%, *n*=145). In contrast, formalization (30%, *n*=62), discovery (26%, *n*=51), and communication (19%, *n*=40) were markedly undervalued; 16% (*n*=33) selected "other". This restricted repertoire of recognized functions may strongly shape how these future teachers conceive proof and proof activity in elementary classrooms.

DISCUSSION, CONCLUSION AND PERSPECTIVES

A large majority of Swiss students selected non-deductive arguments as explanatory: on average of 72.5% of Swiss students (78% for Statement 2 and 67% for Statement 1), which is very close to the 75% in a U.S. sample (Stylianou et al., 2015). Conversely, and again in line with U.S. results, more than half of the Swiss students identified proofs based on deductive reasoning, expressed in a formalized (primarily algebraic) register, as the most rigorous, reflecting a conflation of rigor with formalism. This reveals a persistent tension between the explanatory role of proof and its perceived rigor. Accordingly, future primary teachers in Switzerland recognize (deductive) intellectual proofs as the institutional standard for correctness and rigor, while spontaneously favoring empirical, manipulative, or example-based approaches. At the same time, as in Stylianou et al. (2015), students report limited experience with proof and make largely superficial choices when evaluating proofs. It seems that Swiss students are not fully aware of the limitations of empirical or naïve approaches to proof. In several respects, our results align with earlier studies on secondary students' and future teachers' conceptions of proof reviewed in Stylianou et al. (2015), which consistently report a strong reliance on empirical and example-based approaches. Moreover, this tendency is not accompanied by the use of generic examples that could bridge empirical reasoning and deductive proof. The resulting gap between stated beliefs and mobilized practices points to students' limited engagement with proof during their schooling, particularly the lack of opportunities to produce, discuss, and validate proofs collaboratively. The strong emphasis on verification and explanation, together with minimal recognition of discovery and communication, further reflects a narrow view of proof that overlooks its heuristic and social dimensions. Overall, these findings suggest fragile, context-dependent conceptions of proof shaped by largely passive school experiences, which may hinder future primary teachers' ability to support students' exploratory reasoning and collective mathematical inquiry.

The streamlined design of Parts II and III of the questionnaire appears sufficient to identify dominant student conceptions of proof and their associated difficulties,

yielding results consistent with those obtained through more complex methodologies. This supports the relevance of the instrument as a cost-effective research tool.

Future research should aim to standardize and extend the questionnaire instrument to comparable Francophone populations, particularly prospective primary teachers with similar secondary curricula and classroom practices regarding proof. Such an extension could strengthen its impact on both pre-service and in-service teacher education. More broadly, teacher education should aim to develop prospective primary teachers' conceptions of proof adapted to the mathematical content they will be required to teach, while also emphasizing the social function of proof within a classroom mathematical community. This requires a shift in teachers' professional stance, enabling them to move from naïve or empirical approaches toward rigorous reasoning, while helping them understand that rigor need not depend on high formalism.

References

Balacheff, N. (1988). Aspects of proof in pupils' practice of school mathematics. In D. Pimm (Ed.), *Mathematics, teachers and children* (pp. 216–235). Hodder & Stoughton.

Barkai, R., Tsamir, P., Tirosh, D. & Dreyfus, T. (2002). Proving or refuting arithmetic claims: The case of elementary school teachers. In A. D. Cockburn, E., Nardi (Eds), *Proc. 26th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 2, pp. 57–64). PME.

Gueudet, G. (2008). Investigating the secondary-tertiary transition. *Educational Studies in Mathematics, 67*(3), 237–254.

Hanna, G. (2000). Proof, explanation and exploration: an overview. *Educational Studies in Mathematics, 44*, 5–23.

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

Maher, C. A., Powell, A. B., & Uptegrove, E. B. (Eds.) (2010). *Combinatorics and reasoning: Representing, justifying and building isomorphisms*. Springer.

Ouvrier-Buffet, C. (2023). Exploring students' conceptions of proof in high-school and university: A proposal for collaborations in Europe. In P. Drijvers, C. Csapodi, H. Palmér, K. Gosztonyi, & E. Kónya (Eds.), *Proc. 13th Congress of the European Society for Research in Mathematics Education* (pp. 240–247). ERME.

Stylianides, A.J. (2016). *Proving in the Elementary Mathematics Classroom*. Oxford University Press.

Stylianou, D.A., Blanton, M.L., & Rotou, O. (2015). Undergraduate students' understanding of proof: relationships between proof conceptions, beliefs, and classroom experiences with learning proof. *IJRUME, 1*(1), 91–134.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*, 458–477.