PROOF ASSISTANTS AS CATALYSTS TO DISCUSS ABOUT PROOF IN MATHEMATICS EDUCATION RESEARCH

Carola Manolino[1,4], Antoine Meyer[2,5], Julien Narboux[1,4], Cécile Ouvrier-Buffet[3,6]

[1]Paris Cité University, [2]Gustave Eiffel University, [3]Paris-Est Creteil University, [4]IRIF, [5]LIGM, [6]LDAR – France

This new working group addresses a persistent deadlock in mathematics education research on proof, as identified by Balacheff (2002) and Reid (2005), a challenge that remains still relevant today. Despite an extensive literature, learning and teaching difficulties with proof and proving are worldwide still present, particularly during the secondary-tertiary transition. Multiple theoretical frameworks and experiments – rooted in distinct epistemological, logical, cultural, and cognitive foundations – coexist, often without explicit researchers' positioning. This impedes shared understandings and the collective search for solutions to the didactic challenge of proof instruction. Proof assistants (PAs, e.g., Lean, Rocq etc.) offer a novel lens to revitalize the debate. They are increasingly present in mathematics and computer science research and are gradually implemented in educational contexts (e.g. Hanna et al., 2019) even if not initially designed for educational purposes. While all PAs share the same function of mechanically verifying proofs through formal logic, their interface designs and feedback mechanisms differ substantially. This working group uses their potentials via an unplugged simulation with a shared language to highlight the implicit elements embedded within each researcher's proof conception. It also identifies priority research directions for enhanced community dialogue within mathematics education but also between mathematics education and other domains concern by proof (e.g. computer science, represented by two members of the working group team).

THEORETICAL BACKGROUND

This working group intentionally avoids any predetermined theoretical framework to preserve openness in the debate. Participants are thus encouraged to draw upon the full spectrum of established frameworks in mathematics education research about proof, including Balacheff, Hanna, Duval, Reid, Harel & Sowder, Stylianides and others, for interpreting contrasts emerging from exemplar and PA-simulated proofs.

PLANNED ACTIVITIES

Session 1: from "exemplary proofs" to PA constraints (90 minutes)

The plan is to debate on what a "good" proof is, so to list common ground and differences depending on researchers' points of view. To achieve this aim, following a previous experience (Reid, 2005), participants pre-submit or bring their own most

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: General Contributions*, pp. 330-331. PME.

"exemplary proof" from their research or teaching (on one page). The team of the working group presents the aims and the modalities of the sessions (10 minutes).

Phase 1: Traditional mode (40 minutes)

In small groups of four to five, participants share exemplars. Each participant controls at least one colleague's proof. In each group, the participants share the justifications of their choices and their control criteria. These are synthesized in plenary.

Phase 2: What is a PA? For what? (40 minutes)

What a PA is and how it works will be illustrated by the working group team with two different PAs (a point-and-click PA and a controlled natural language PA). The *tactics* and validation processes are highlighted.

Session 2: Simulated assistant mode and debrief (2×45 minutes)

Participants rework their "exemplary proof" under an unplugged PA, with constraints provided by the working group team (limited tactics to rework the proof and simulated automated validation). Groups document the way they have to adapt or modify their reasoning processes, their proof structure, the granularity of their arguments, the language. Then, they analyze the contrasts observed between the two modes of proving (traditional mode and simulated assistant mode). They identify dimensions along which the two proofs differ, they articulate how validation operates and they examine whether the simulated mediation changes what counts as a "satisfactory" proof. Each group produces a written synthesis, identifying key analytical criteria emerging from the comparison and concrete research questions suggested by these observations. A final plenary discussion summarizes contributions into a shared set of analytical dimensions and research questions about the epistemology of proof and the use of PAs in teaching and learning.

EXPECTED OUTCOMES

The expected outcome is to produce a first shared analytical grid derived from exemplary proofs / PA contrasts for PME newsletter. A follow-up activity includes a working group at PME50 to continue the grid and the debate about the relevance of PAs in teaching and learning proof.

References

Balacheff, N. (2002). The researcher epistemology: a deadlock from educational research on proof. In Fou Lai Lin (Ed.), *International Conference on Mathematics - "Understanding proving and proving to understand"* (pp. 23-44). NSC and NTNU.

Hanna, G., Reid, D.A., De Villiers, M. (2019)(Eds.). *Proof technology in mathematics research and teaching.* Springer.

Reid, D. (2005). The meaning of proof in mathematics education. In M. Bosch (Ed.), *Proceedings of the 4th Conference of the European Society for Research in Mathematics Education* (pp. 458-468). ERME.