My Take: Proof Research at the Undergraduate Level--How it Has Evolved[1]

Annie Selden, New Mexico State University[2]

1. Early Days: The Beginnings of Mathematics Education Research on Proof at the Undergraduate Level

Although there had been mathematics education research on proof more generally (e.g., Bell, 1976; Fischbein, 1982), there was little mathematics education research on proof at the undergraduate level until about the mid-1980s. John and my first foray into such research began in the mid-70s when we were teaching at the University of the Bosphorus in Istanbul, Turkey, and submitted an article on logical reasoning errors that students made in an abstract algebra course taught Moore Method, published in a local journal (Selden & Selden, 1976). This article was later recast in terms of misconceptions research and presented at a Cornell University conference (Selden & Selden, 1987), and is sometimes cited today. At about the same time, unbeknownst to us, Ed Dubinsky was researching topics on undergraduates' understandings of logic (Dubinsky, 1987, 1989; Dubinsky, et al., 1988). As far as I can tell, many of these early empirical publications were based on careful observations of teaching undergraduates. Research on, and interest in undergraduate students' knowledge of, and use of, logic in proving has continued to the present time (e.g., Dawkins & Roh, 2022; Dawkins & Norton, 2022; Durand-Guerrier, et al., 2012, Savic, 2012).

At about the same time, Gila Hanna, whose mathematics education research has often focused on the more theoretical and philosophical aspects of proof and proving, made the now well-known and influential distinction between "proofs that only prove and proofs that explain" (Hanna, 1989, 1990). Also, Michael de Villiers, whose mathematics education research has often focused on the teaching and learning of geometry, proposed five functions of proof which are still quoted today; namely, proof as means of: (1) *verification/conviction*; (2) *explanation*; (3) *systematization*; (4) *discovery*; and (5) *communication* (de Villiers, 1990). De Villiers stated that this analysis was based on epistemological considerations and personal testimonies of practicing mathematicians. Still, a reading of the article suggests it was also based on insightful observations of his teaching and reading of the available literature (e.g., Alibert, 1988; Freudenthal, 1973; Hanna, 1989; Lakatos, 1976; Wilder, 1944).

[1] This is a preprint of a chapter for a book, being edited by Keith Weber and Milos Savic, anticipated to be published Summer 2023, and tentatively titled, *New Directions in University Proving: Honoring the Legacy of John and Annie Selden.*

[2]Dr. Annie Selden is Professor Emerita, Tennessee Technological University. After retirement in 2003, she and her husband, Dr. John Selden, took up adjunct positions in the Department of Mathematical Sciences, New Mexico State University, where they directed four PhD students in research in undergraduate mathematics education, including Milos Savic.

Another early foray into mathematics education research at the undergraduate level came with the formation in 1985 of the Advanced Mathematical Thinking Working Group of the International Group for the Psychology of Mathematics Education (PME). The idea behind this Working Group was to focus on "advanced mathematical thinking", in contrast to much of the prior work of PME, which had concentrated on "elementary mathematical thinking" (Harel, Selden, & Selden, 2006, p. 147). Because the members of this Working Group came from a variety of countries, for convenience, it was decided that the term "advanced mathematical thinking" would be used for education beyond the compulsory stage, which at that time often concluded at age 16 and certainly included mathematics education research at the undergraduate level. A major undertaking of this Working Group was the writing of a book on advanced mathematical thinking (Tall, 1991); relevant to early research on proof, there were book chapters specifically on proof (Hanna; Alibert & Thomas), on the role of definitions, including a short discussion of concept image and concept definition (Vinner), and on research in the teaching and learning of mathematics at an advanced level (Robert & Schwarenberger). Other researchers have since provided their own definitions/descriptions of advanced mathematical thinking (e.g., in a special issue of *Mathematical Thinking and Learning* edited by Selden & Selden, 2005).

1.1 Experiencing Proof and Proving in the Classroom, the Method of Scientific Debate, and Structured Proofs

While not about proof at the undergraduate level, Balacheff (1988), in an empirical study of French secondary pupils' proving, categorized their arguments into four different types and argued that these represented four increasingly sophisticated levels of thinking: (1) *naïve empiricism*, in which an individual arrives at a conclusion on the validity of an assertion based on a small number of confirming cases; (2) *crucial experiment*, in which an individual considers the possibility of generalization by examining a case that is not very specific; (3) *generic example*; in which the proof rests upon properties which are a generalization of a class of examples; and (4) *thought experiment*, in which an individual can distance themselves from specific actions and make logical deductions based only on knowledge of the properties and relationships characteristic of the situation. These categorizations are sometimes still referred to today (e.g., Varghese, 2011).

The work of Alibert and Thomas (1991) on the *method of scientific debate* in French university classrooms was an attempt to get undergraduates to understand proofs, because they had observed that the students had difficulty understanding proofs when they read through textbook proofs in their strict formal, linear order. They wanted to "enable students to see proof as a necessary part of the scientific process of advancing knowledge, rather than just a formal exercise to be done for the teacher," (Alibert & Thomas, 1991, p. 224). They described three steps in generating a scientific debate: (1) The teacher initiates and organizes the production of scientific statements [conjectures] by the students. (2) The students then provide support for the conjecture, or not, by scientific argument, proof, refutation, or counter-example. (3) The statements that can be confirmed by a full demonstration become theorems accepted by the class. (Alibert & Thomas, 1991, p. 225).

Another attempt to get away from the strict linear order of presenting proofs to university mathematics students was a structural method suggested by Uri Leron (1983). He wrote

> The method, triggered by recent ideas from computer science, is intended to increase the comprehensibility of mathematical presentations while retaining their rigor. The basic idea underlying the structural method is to arrange the proof in levels, proceeding from the top down; the levels themselves consist of short autonomous "modules," each embodying one major idea of the proof. (Leron, 1983, p. 174).

While this structural presentation method was long thought to be a plausible one, it was "put to the test" empirically much later in a qualitative study that presented structured proofs to university mathematics students to see how they "read and perceived this type of proof presentation. Although some students valued the summaries contained in structured proofs, many complained that structured proofs 'jumped around' and required them to scan different parts of the proof to coordinate information." (Fuller, et al., 2014, p. 1).

2. Into the 1990s--Further Developments in Mathematics Education Research on Proof at the Undergraduate Level

The 1990s saw a gradual flowering of mathematics education research at the undergraduate level. While much of this research concentrated on the teaching/learning of individual subjects (e.g., calculus, see Selden, Mason, & Selden, 1989; linear algebra, see Dorier, 1998) or on mathematical concept acquisition more generally (e.g., functions, see Dubinsky & Harel, 1992), some of it was devoted to proof and proving at the undergraduate level.

2.1 Undergraduate Students' Difficulties in Constructing Proofs

For example, John and I published research on undergraduate students' understanding of, and ability to unpack, logical statements (Selden & Selden, 1995). That study focused on transition-to-proof course students' ability to unpack informally written mathematical statements into the language of predicate calculus. The general notion of *unpacking* "has been applied variously to symbols, graphs, and diagrams. For example, when a group is denoted by G, realizing that a more expansive notation is (G, +) and evoking the group axioms, is an unpacking of the symbol G." (Selden & Selden, 1995, p. 128). For this research, the statement *For a < b, there is a c so that f(c) = y whenever f(a) < y and y <f(b)* could be correctly unpacked as $\forall a\ \forall b\ \forall f\ \forall y\ \exists c\ [(a < b \wedge f(a) < y \wedge y < f(b)) \rightarrow f(c)=y]$. For such "informal calculus statements, just 8.5% of unpacking attempts were successful; for actual statements from calculus texts, this dropped to 5%." We inferred that "these students would be unable to reliably relate informally stated theorems with the top-level logical structure of their proofs and hence could not be expected to construct proofs or validate them, i.e., determine their correctness." (Selden & Selden, 1995, p. 123).

Like much of the research on proof and proving in this time period, our research focused on what undergraduate students couldn't do (Selden & Selden, 1987, 1995). Perhaps the underlying motivation for doing such research was the idea that research should obtain some baseline information on what students were currently learning, and could do in the way of proving and problem solving, with then current university teaching. For example, Robert Moore did his dissertation research at the University of Georgia on transition-to-proof course students' understandings of formal proof. He found that "An inductive analysis of the data revealed three major sources of the students' difficulties: (a) concept understanding, (b) mathematical language and notation, and (c) getting started on a proof." (Moore, 1994, p. 249). Somewhat later in the decade, Keith Weber conducted his dissertation study at Carnegie-Mellon University on undergraduate and doctoral abstract algebra students' proving—he documented that the undergraduates in the study were unable to apply facts they knew to prove theorems on groups. He hypothesized that the undergraduates failed "to construct a proof because they could not use the syntactic knowledge that they had." In contrast, the doctoral students "appeared to know the powerful proof techniques in abstract algebra, which theorems are most important, when particular facts and theorems are likely to be useful, and when one should or should not try and prove theorems using symbol manipulation." (Weber, 2001, p. 101).

2.2 Undergraduate Students' Proof Schemes

There was also work in the 1990s on undergraduate students' ideas of proof and proving, which included work on how they construct proofs and what they see as constituting a proof. In particular, Harel & Sowder (1998) classified students' *proof schemes*. By *proving* they meant "*the process employed by an individual to remove or create doubts about the truth of an observation* [conjecture]." They described the process of proving as consisting of two subprocesess. (1) "*Ascertaining is the process employed by an individual to remove or create doubts about the truth of an observation* [conjecture]." (2) "*Persuading is the process an individual employs to remove others' doubts about the truth of an observation* [conjecture]." (Harel & Sowder, 1998, p. 241, italics in the original). They described, and gave examples of, three overarching proof scheme categories: (1) *External conviction proof schemes*--basically it's a proof if someone knowledgeable or a textbook tells you it's a proof; (2) *Empirical proof schemes*—checking enough cases to convince you that the conjecture is true; and (3) *Analytical proof schemes*—broken down further into *Transformational proof schemes* and *Axiomatic proof schemes*—basically proofs that would be acceptable to the mathematical community. The notion of proof schemes has been widely used to analyze what constitutes for proof students (e.g., Erikson & Lockwood, 2021; Housman & Porter, 2003).

2.3 Establishment of RUME Organizations

By 1998, there was enough research on the teaching and learning of university mathematics to convene an ICMI Study Conference in Singapore, with the actual book published a bit later (Holton, 2003). While there are book chapters on research into the teaching/learning of calculus and linear algebra, on the secondary-tertiary transition, on APOS (Action, Process, Object, Schema) theory, and 51 entries on proof in the index, no single chapter was devoted to proof at the undergraduate level, suggesting that the organizers did not yet see, or know about, significant research on proof and proving at the undergraduate level.

However, in the USA, the beginnings of research in undergraduate mathematics education (RUME) were emerging, as indicated by the three conferences on RUME in 1996, '97, and '98, organized by the RUMEC (Research in Undergraduate Mathematics Education Community) group under the leadership of Ed Dubinsky. This was followed in 2001 by the formation of the first Special Interest Group of the Mathematical Association of America on Research on Undergraduate Mathematics Education (SIGMAA on RUME). The SIGMAA on RUME continues the tradition of holding annual RUME Conferences, at which there are always presentations on proof, proving, or proof comprehension (e.g., Moore, Byrne, Hanusch, & Fukawa-Connelly, 2016).

3. In the Early 2000s, Mathematics Education Research on Proof Started to Broaden to Consider Topics Such as Proof Validation, the Teaching of Proof, Affect During Proving, and Proof Comprehension

3.1 Early Validation Studies

Our early exploratory study of eight transition-to-proof course students' ability to validate, that is, to determine the correctness of, proofs investigated how these students "read and reflected on four student-generated arguments purported to be proofs of a single [elementary number theory] theorem." (Selden & Selden, 2003, p. 4). This was not our very first foray into validation--we had provided a sample, hypothetical proof validation of a calculus theorem in Appendix 1 of our "unpacking paper" (Selden & Selden, 1995, pp. 143-147). In this study of transition-to-proof course students' ability to correctly validate other similar students' proof attempts, we found it to be at chance level. Subsequently, a number of other mathematics education researchers took up proof validation studies. Building on this research, further studies examined how mathematicians validate proofs (Weber, 2008) and how mathematicians read proofs (Weber & Mejia-Ramos, 2011). Even later, proof validation research was subsumed by some under the more general heading of proof comprehension research—how individuals read, comprehend, and learn from presented proofs. However, validating a proof attempt, either one's own or another person's, seems to require somewhat different skills than comprehending a presented proof, whether in a text or a lecture, although there is some overlap. (Selden & Selden, 2015, p. 343).

3.2 Early Research on the Teaching of Proof at the Undergraduate Level

How proof is actually taught at the undergraduate level, while not much researched up to this time, was investigated by Keith Weber (2004) in his study of the entirety of one professor's introductory real analysis course. He identified three separate teaching styles that the professor used: (1) a *logico-structural teaching style* in the case of sets and functions; (2) a *procedural teaching style* in the case of limits of sequences; and (3) a *semantic teaching style* in the case of elementary topological concepts like interior point of a set. This study revealed, amongst other things, that the popular, stereotypical idea that the teaching of university advanced mathematics courses, such as real analysis, consisted "entirely of definition, theorem, proof, definition, theorem, proof, in solemn and unrelieved concatenation" (Davis & Hersh, 1981, p. 151) was too narrow and needed further investigation.

Additional studies of actual university mathematics teaching were not taken up until after it was pointed out that the teaching of proof-based university mathematics courses was in need of investigation (Speer, Smith, & Horvath, 2010). A recent literature review of 104 published papers reporting research on the teaching of proof-based mathematics courses at university considered both lecture-based and student-centered pedagogies. For each type of instruction, the authors described the instruction, instructor beliefs and rationales, and the relationship between instruction and students' learning. (Melhuish, Fukawa-Connelly, Dawkins, Woods, & Weber, 2022). The authors noted that often the studies were hard to compare due to the use of different theoretical frameworks. They also observed that there are still too few studies today that attempt to link instructors' teaching of proof to university students' learning of proof.

3.3 The Beginnings of Incorporating Affect into Research on Undergraduates' Proving

In this decade, we began thinking about how affect, in particular consciousness and non-emotional cognitive feelings, might be involved in the proving process (e.g., Selden, McKee, & Selden, 2008, 2010). Previously affect had often been viewed as separate from, but related to cognition; affect had also been divided into considerations of *beliefs*, *attitudes*, and *emotions*. These were described as being of increasing intensity and decreasing stability, with emotions the most intense (McLeod, 1989). At this time, we were not the only ones who were thinking, and writing about the role of affect and its relation to cognition; for example, DeBelllis and Goldin (2006) considered the relation of affect to mathematical problem solving, adding a fourth component, namely, *values*.

3.3.1 Our Take on Affect in Proving

In our paper, we focused on "a particular kind of affect – *nonemotional cognitive feelings* – and on the implementation of actions via behavioural schemas." (Selden, McKee, & Selden, 2010, p. 199). Some examples of such nonemotional cognitive feelings are "feelings of knowing, of caution, of familiarity, of confusion, of not knowing what to do next, of rightness/appropriateness, of rightness/direction or of rightness/summation." (p. 202). Feelings can provide information, which can be positive or negative. Indeed "the feeling in constructing a proof that one is 'on the right track' is a cognitive feeling." (p. 203). Such feelings can lead to actions via *behavioural schemas*, a notion for which we provided a six-point theoretical sketch:

> (1) Behavioural schemas are immediately available. They do not normally have to be searched for, which distinguishes them from most conceptual knowledge and episodic and declarative memory.
>
> (2) Simple behavioural schemas operate outside of consciousness. One is not aware of doing anything immediately prior to the resulting action – one just does it. They are not under conscious control.
>
> (3) Behavioural schemas tend to produce immediate action--one becomes conscious of the action as it occurs or immediately after it occurs
>
> (4) Behavioural schemas cannot 'chained together' – they function entirely outside of consciousness and one is consciousness of only the final action.
>
> (5) An action due to a behavioural schema depends on conscious input, at least in large part.

> (6) Behavioural schemas are acquired through, possibly tacit, practice, that is, to acquire a beneficial schema a person should actually carry it a number of times – not just understand its appropriateness. Changing a detrimental behavioural schema requires similar, perhaps longer, practice. (This is a summary of Selden, McKee, & Selden, 2010, pp. 205-206).

Following this description, we illustrated a number of actual examples of behavioural schemas that we had observed in undergraduate proving situations.

While there has been much research on what one might call "hot" emotions, such as mathematical anxiety (e.g., Rozgonjuk, et al., 2020), there has been less research on more useful and calm emotions. One recent example of such research concerns undergraduate transition-to-proof course students' satisfying moments, including "understanding, overcoming challenges, and accomplishments without struggle". (Satyam, 2020). However, in my view, there is a need for more research on the role of feelings, and affect more generally, in proving.

Also, while there has been a great deal of international interest on affect in mathematics education research (e.g., see the ICME-13 monograph edited by Hannula, et al., 2019), most research deals with affect during problem solving, rather than proving, although constructing a proof can be seen as a kind of problem solving.

3.4 Observing that Proof Comprehension is Under Researched

At the 19th ICMI Study Conference on Proof and Proving in 2009, Juan Pablo Mejia-Ramos and Matthew Inglis presented the results of a bibliographic study of the "different argumentative activities associated with the notion of mathematical proof" that had been done up to that time. In their sample of 131 empirical research articles (pared down from an original 641), they found that of

> those articles in our sample that discussed specific tasks, the majority (82 papers) addressed students' construction of novel arguments, some (24 papers) involved students' reading of given arguments and none focused on the presentation of a given argument. In particular, only 3 articles addressed tasks related to the comprehension of a given argument and none of the articles discussed tasks directly focussed on the presentation of an argument to demonstrate students' understanding of it. (Mejia-Ramos & Inglis, 2009, p. 91).

This observation subsequently led to a burgeoning of research on proof comprehension in the next decade, beginning around 2010, much of it done by the Proof Comprehension Research Group at Rutgers University.

3.5 An Assessment Model for Proof Comprehension Research

Having established that proof comprehension was under researched (Mejia-Ramos & Inglis, 2009), Juan Pablo Meija-Ramos and others embarked on proof comprehension research, beginning with the development of a multi-dimensional assessment model for proof comprehension at the undergraduate level (Mejia-Ramos, Fuller, Weber, Rhoads, & Samkoff, 2012). The authors contended that

> in undergraduate mathematics a proof is not only understood in terms of the meaning, logical status, and logical chaining of its statements but also in terms of the proof's high-level ideas, its main components or modules, the methods it employs, and how it relates to specific examples. (p. 1).

The authors maintained that, to demonstrate comprehension of a proof, students ought to be able to understand the meaning of terms in the proof by being able to:

(1) State the definitions of terms used in the proof in their own words. (2) Identify trivial implications of a given statement. (3) Identify examples that illustrate a given term in the proof. (4) Restate a given statement in a different, but equivalent, manner.

Also, students should be able to understand the logical structure and proof framework by being able to:

> (5) Identify the type of proof framework (e.g., contradiction, contraposition). (6). Identify the purpose of a sentence within the proof framework.

In addition, students should be able to justify claims in the proof by being able to:

> (7) Make explicit an implicit warrant in the proof. (8) Identify specific data supporting a given claim. (9) Identify the specific claims that are supported by a given statement by answering questions such as: Which claims in the proof logically depend on a given line of the proof?

Furthermore, students should understand the overall structure of the proof by being able to:

> (10) Provide a good summary of the proof. (11) Identify a good summary of a key sub-proof.

Moreover, students should be able to identify the modular structure of the proof, by being able to:

> (12). Partition the proof into modules. (13). Identify the purpose of a module in the proof. (14). Identify the logical relation between modules of the proof.

Finally, students should be able to transfer the general ideas or methods of the proof to another context, by being able to:

> (15) Identify the method of proof. (16) Transfer the method of proof to a different task. (17) Appreciate the scope of the method. (This is a summary and renumbering of Mejia-Ramos, Fuller, Weber, Rhoads, & Samkoff, 2012, pp. 5-19).

How the above seventeen abilities might be demonstrated by an individual is illustrated via an analysis of the proof of the number theory theorem: *There exist infinitely many triadic primes.* One can wonder whether very many undergraduate, or even beginning graduate, students could demonstrate all these abilities, even when specifically asked to. However, more recently, there has been work on the development and validation of proof comprehension tests (e.g., Mejia-Ramos, Lew, de la Torre, & Weber, 2017).

3.6 Some Theoretical Ideas Considered in the First Decade of the 2000s: *Semantic and Syntactic Proof Productions, Key Ideas of a Proof, Cognitive Unity, Toulmin's Scheme*

Several distinctions regarding students' proving were made in this decade (2000-2010). For example, Weber and Alcock (2004) distinguished

> two ways an individual can construct a formal proof. We define a *syntactic proof production* to occur when the prover draws inferences by manipulating symbolic formulae in a logically permissible way. We define a *semantic proof production* to occur when the prover uses instantiations of mathematical concepts to guide the formal inferences that he or she draws. (p. 239).

They presented exploratory case studies from undergraduate group theory and real analysis proof attempts. While the idea of contrasting semantic and syntactic proof productions was taken up again (Mejia-Ramos, Weber, & Fuller, 2015), this binary distinction between syntactic and semantic proof productions seems too stark; proving is more likely to be a combination of the two ways of arguing/reasoning. In a recent article, Weber (2021) considered the role of syntactic representations in set theory, where this distinction seems more applicable.

Another binary distinction was put forward by Manya Raman (2003) who introduced the notion of *key idea(s)* of a proof. She developed a framework "for characterizing people's views of proof, based on a [binary] distinction between public and private aspects of proof and the key ideas which link these two domains." (Raman, 2003, p. 319). Since being introduced, the notion of key ideas has sometimes been considered further (e.g., Yan & Hanna, 2019). Also, around this time, another binary distinction was introduced by Paolo Boero and colleagues (1996) -- the idea of the *cognitive unity* between the processes of conjecturing and proving, although it can be difficult to observe and analyze the continuity of these two processes. (Pedemonte, 2007, p. 25), and again this binary distinction is too stark.

For a while, there was some interest in using Toulmin's (1958) argumentation scheme, developed to analyze philosophical and legal arguments, to analyze proofs and proving. However, the concepts of claim, data, warrant, backing, and especially the concepts of modality and rebuttal, seem more suited to analyzing classroom argumentation and conjecturing (e.g., Fukawa-Connelly, 2014), than to proof construction.

4. Research on Proof at the Undergraduate Level Expands in a Variety of Directions After About 2010 to the Present[i]

The number of researchers investigating proof at the undergraduate level and the variety of topics investigated now expanded, one might say almost exploded, in a multitude of directions: further research on proof construction, validation, and comprehension; on the university teaching of proof and proving and proof-based courses; on mathematicians' reading, sometimes skimming, of proofs; on expert/novice studies of proof comprehension, including eye-tracking studies; on investigation of what university students learn from lecture and inquiry-based courses; and on mathematicians' and university students' knowledge of, and valuing of, the genre of proof, to name but a few.

4.1 Studies of Mathematicians' Teaching of Proof and Proving: Their Views of Successful Students' Thinking, Their Classroom Teaching, How They Grade, IBL teaching

Mathematicians have views, sometimes strong ones, on how to teach proof-based courses. One can wonder: Where did they get their views on the teaching of proof/proving? Probably from their own experiences, taking undergraduate and graduate proof-based mathematics courses. Almost surely, most have not read the mathematics education research literature, such as it is, regarding what is now known about effective, and ineffective, university mathematics teaching, although more recently some information from the research literature has appeared in publications that mathematicians read, such as the *Notices of the AMS* (e.g., Alcock, Hobbs, Roy, & Inglis, 2015).

4.1.1 Mathematicians' views of successful provers and studies of how mathematicians teach

In an exploratory interview study of five mathematicians' views on their teaching of a course meant to introduce students to mathematical reasoning and proof, Lara Alcock (2010) identified "four modes of thinking that these professors indicated are used by successful [university student] provers." (p. 73). These are: (1) *instantiation*, meaning understanding a mathematical statement by thinking about particular or generic objects to which it applies; (2) *structural thinking*, meaning generating a proof for a statement by using its formal structure, that is, making formal deductions based on the statement and/or associated definitions and known results; (3) *creative thinking*, meaning examining instantiations of mathematical objects in order to identify a property or set of manipulations that can form the crux of a proof; and (4) *critical thinking*, meaning checking the correctness of assertions by looking for counterexamples and properties that are implied or should be preserved.

In a study based on observations of one abstract algebra instructor's lecture-based teaching, Tim Fukawa-Connelly (2012) documented that "she [the instructor] frequently modeled the aspects of hierarchical structure and formal–rhetorical skills, and structural, critical, and instantiation modes of thought" (p. 325). She also attempted to involve students by asking questions, but most required only a factual response. The instructor was attempting to model the way an expert in the discipline thinks.

A separate, later video case study of proofs presented in one real analysis professor's (Dr. A's) lectures, investigated why students did not understand what the professor wanted to convey. Dr. A was justifying the claim: *If a sequence* $\{x_n\}$ *has the property that there exists a constant r with* $0 < r < 1$ *such that* $|x_n - x_{n-1}| < r_n$ *for any two consecutive terms in the sequence, then* $\{x_n\}$ *is convergent.* Dr. A was interviewed as to what he intended to convey and six students were also asked to describe what were the main ideas of the proof while looking at their notes and twice at a video of the classroom proof. Dr. A noted that he had five ideas about Cauchy sequences that he wished to convey, four of which were useful methodological ideas, that he had described informally by using words like *small*, but did not write on the blackboard. It turned out that the students "did not grasp many of the ideas that Dr. A emphasized as the most important parts of the lecture, even after viewing the videos of the lecture for a second time." (Lew, Fukawa-Connelly, Mejia-Ramos, & Weber, 2016, p. 185).

One wonders how best to advise students that it might be better not to copy the proof verbatim from the blackboard, but rather to take notes on what the professor says about the proof and its importance. Perhaps one way might be for instructors to use Dr. T's idea of handing out a hard copy of the complete proof prior to presenting/discussing a proof, so students could concentrate on more informal explanations and diagrams (Weber, 2001, p. 127), but this would need to be tested.

4.1.2 Pedagogical proofs and why some proofs seem hard for students

Mathematicians' views of what they consider to be good pedagogical proofs has been investigated. In a qualitative study of how eight mathematicians revised two analysis proofs for presentation to mathematics majors, Lai, Weber, and Mejia-Ramos (2012) found that the mathematicians thought "that introductory and concluding sentences should be included in the proofs, main ideas should be formatted to emphasize their importance, and extraneous or redundant information should be removed to avoid distracting or confusing the reader." In a second larger quantitative study (N=110) to assess whether other university mathematics instructors agreed with the eight mathematicians of the first study, it was found that there was "a high degree of agreement among mathematicians regarding how they would revise proofs for pedagogical purposes." (p. 146).

While it is known that some types of proofs seem harder to construct or comprehend (e.g., contradiction proofs, mathematical induction proofs, recipe proofs), is there some way, other than using the logical order of course topics, to design the teaching of proof, for example, in a transition-to-proof course? There were earlier studies of students' difficulties with proof by contradiction (e.g., Antonini & Mariotti, 2006). More recently Rabin & Quarfoot (2021) reported that, for their transition-to-proof course participants, "the knowledge resources students bring to bear on proof problems, and how these resources are activated, explain more of their 'difficulties' than does the logical structure of the proof technique."

4.1.2.1 Proofs by Mathematical Induction

Mathematical induction proofs have long been considered difficult for students to understand and construct. When Gila Hanna (1990) introduced the notion of "proofs that explain versus proofs that only prove", she used the fact that the sum of the first n integers is $n(n+1)/2$, when proved formally by the Principle of Mathematical Induction (PMI), as a "proof that only proves".

To help students with PMI, Harel (2002) reported on a "fundamentally different instructional treatment of mathematical induction", in which prospective secondary teachers' conception of mathematical induction developed as a transformational proof scheme. The treatment was guided by a system of learning-teaching principles, called the DNR system, which is an acronym for the principles of Duality, Necessity, and Repeated Reasoning, developed previously by Harel (1998). The *Duality Principle* makes the distinction between ways of thinking and ways of understanding, which are somewhat different from what one might suppose (see the editors' comments on Harel, 2008). The *Necessity Principle* states that students are likely to learn if they see a need for what we teach them, by which Harel meant an intellectual need as opposed to a social or economic need. The *Repeated Reasoning Principle* states that students must practice reasoning in order to internalize and interiorized specific ways of thinking and ways of understanding.

More recently Relaford-Doyle and Núñez (2021), "used a 'visual proof by induction' – a simple image that is designed to demonstrate a theorem that would be formally proven using mathematical induction – to investigate students' conceptualizations of mathematical induction." They found that

> First, the majority of students who were familiar with formal mathematical induction had difficulty using the image to justify the theorem, suggesting that their knowledge of the proof method was intimately linked to the algebraic method and thus largely procedural in nature. Second, students who had not studied formal mathematical induction generally used the image as the basis of a standard inductive generalization and did not recognize that the image could be used to establish the necessity of the theorem. Surprisingly, these students often expressed conceptualizations of natural number that were inconsistent with the formal characterization that forms the basis of formal mathematical induction. (Relaford-Doyle & Núñez, 2021, p. 1).

4.1.3 Mathematicians' grading of student proof submissions

Students not only learn proving from their classroom experiences, whether lecture or inquiry-based, they learn from how their proofs are scored, and especially, from their instructors' comments, provided they read them. Moore (2016) conducted an exploratory study of four mathematicians' evaluation and scoring of six undergraduate student proofs in a discrete mathematics or geometry course and asked them questions about the characteristics of a well-written proof and how they communicated these to students. They agreed the

> most important characteristics of a well-written proof are logical correctness, clarity, fluency, and demonstration of understanding … [and although] the professors differed in the attention they gave to fluency issues, such as mathematical notation, layout, grammar, and punctuation, they agreed in giving these characteristics little weight in the overall score. (p. 246).

In a similar study, nine mathematicians were interviewed as they "assigned points to three student-generated [number theory] proofs from a transition-to-proof course." Curiously, there were

> ten instances in which a mathematician did not assign full credit to a proof that they evaluated as correct [primarily due to minor omissions] … [and similar to Moore (2016)] mathematicians assigned points based not primarily on the correctness of the written artifact that they were given, but rather based on their models of students' understanding." (Miller, Engelke, & Weber, 2018).

One can wonder what messages students are getting about the genre of proof—what's important to include, how much explanation/justification is needed, and so forth.

4.1.4 Research on IBL teaching

The phrase, Inquiry Based Learning (IBL), has been used in the mathematics community since the 1990s by those, mainly US, mathematicians attracted by aspects of the Moore Method (cf., Coppin, et al., 2009). More recently, the mathematics education research community has taken an interest in understanding the methods of, and potential of, IBL teaching (e.g., Haberle, et al., 2018).

> [W]hile it did not originate from educational research, its practices are generally well aligned with research and there are good pedagogical foundations for an inquiry-based approach to learning. What makes it special is its development as a social phenomenon, bringing together those who have discovered the power of student-centered and active learning approaches to undergraduate mathematics instruction. It now embraces a "big tent" that is characterized by a willingness to cede large chunks of time from lecture so that students have structured opportunities within class to explore concepts, investigate key ideas, and build understanding through inquiry. (Bressoud, 2019).

IBL has similarities with Inquiry-Oriented Instruction (IOI), which was developed within the mathematics education research community, while IBL did not.

> While IBL has become a social movement for a wide variety of practitioners of student-centered learning, IOI is strongly rooted in the discipline of research in mathematics education, drawing directly on educational theory to explore student thinking and build activities that seek to address common difficulties. (Bressoud, 2019).

In my view, IBL is more often focused on developing university students' proving skills[ii], whereas IOI is often more concentrated on having students develop mathematical concepts in such topics as beginning differential equations (e.g., Rasmussen & Kwon, 2007), linear algebra, and group theory. (cf., Johnson, et al., 2013).

4.2 Research on Mathematicians: *Experts' Versus Novices' Reading of Proofs*; *Eye-Tracking Studies, Proof Summaries, Explanatory Value of Proofs, Use of Examples in Proving, Genre of Proof*

Researchers have investigated what mathematicians do when they read and construct proofs. This seems to parallel, but probably was not influenced by, psychological research on expert versus novice behavior (e.g., Schunn & Nelson, 2009). The experts in the mathematics education research studies are mathematicians whose behavior and views, with regard to proof and proving, are investigated.

4.2.1 Experts' versus novices' reading of proofs

A more recently used methodology to gain information on mathematicians' and students' reading of proofs is eye-tracking, which allows researchers to investigate what individuals focus on and how they traverse the written text. Inglis and Alcock (2012) in such an eye-tracking study, found

> compared with mathematicians, undergraduate students spend proportionately more time focusing on “surface features” of arguments, suggesting that they attend less to logical structure; and … compared with undergraduates, mathematicians are more inclined to shift their attention back and forth between consecutive lines of purported proofs, suggesting that they devote more effort to inferring implicit warrants. (Inglis & Alcock, 2012, p. 358).

Other researchers have taken up similar eye-tracking studies. For example, in one rather specialized study of eight mathematicians’ reading of proofs having accompanying pictures, it was found that all paid attention to the pictures, but “in two out of three items, the text was fixated upon significantly longer than the picture. The data suggest that the participants tried to integrate information from text and picture by alternating between these representations.” (Beitlich, et al., 2014).

4.3.2 Proof summaries

Because writing good proof summaries is one indication of proof comprehension [see (9) in Section 3.5 above], a recent study investigated features of students’ proof summaries that were most valued by expert judges [mathematicians]. In that study, students were first provided “with a proof that the open unit interval is uncountable, and asked for a 40-word summary.” Subsequently, mathematicians were asked to make comparative judgments[iii]. It was “found that high-scoring summaries referenced a proof’s logical structure and the mechanism by which it reached a contradiction.” (Davies, et al., 2020, p. 181). Perhaps it would also be good to investigate how university students might be taught to write good proof summaries. On the web, one can find tips for writing good summaries of more general text, but it is not clear these would apply to proof summaries.

4.3.3 Explanatory value of proofs

In the more philosophical literature, “proofs that explain” have often been valued over “proofs that only prove”, with visual proofs considered to be explanatory. However a recent empirical study of mathematicians’ views found that “in contrast to claims made in the literature regarding the explanatory value of different types of proofs, mathematicians in our study did not seem to judge visual proofs as particularly explanatory” (Mejia-Ramos, Evans, Rittberg, & Inglis, 2021). As to mathematicians’ notion of explanatoriness, and its relationship to prior accounts of mathematical explanation (e.g., Hanna & Jahnke, 1996), the authors further stated that

> Using a Comparative Judgement approach, we asked 38 mathematicians to assess the explanatory value of several proofs of the same proposition. We found an extremely high level of agreement among mathematicians, and some inconsistencies between their assessments and claims in the literature regarding the explanatoriness of certain types of proofs. (Mejia-Ramos, Evans, Rittberg, & Inglis, 2021, p. 575).

4.3.4 Use of examples in proving

An early discussion of *generic examples* was given by Mason and Pimm (1984), who observed that “A generic example is … presented in such a way as to bring out its intended role as the carrier of the general. This is done by stressing and ignoring various key features." (p. 287). Somewhat later Balacheff (1988; see Section 1.1) observed students using generic examples in proving.

Generic proofs are based on generic examples and have been used in university mathematics teaching. Rowland (2002) gives a number of nice examples of generic proofs that he uses in undergraduate number theory and suggests principles for how to select them. These include: (1) The particular case should be neither too trivial nor too complicated. (2) In number theory, where the proofs are often about primes, one should pick a number like 17 that one can easily follow through in a generic proof. (Rowland, 2002, pp.167-168).

More recently, a study of successful mathematicians' and students' use of examples, in exploring and proving conjectures, observed how participants selected and used them (Ellis, et al., 2013). Reasons for example choice were to: (1) Test boundaries of a hypothesis; (2) Examine relevant properties to the conjecture; and (3) Build a progression of specific examples. Participants used examples to: (1) Attend to common features across multiple examples; (2) Identify the structure of a general argument; and (3) Envision an example as a changing representation. (This is a summary of Table 2; Ellis, et al., 2013, p. 268).

4.3.5 Genre of proof

We discussed, and gave some examples of, the genre of proof in a paper presented at the Ben Gurion University Symposium in honor of Ted Eisenburg's retirement (Selden & Selden, 2013). Despite initial objections of the organizers regarding our use of the phrase "genre of proof", it has since been taken up by others (e.g., Bowers & Küchle, 2020). Also, unbeknownst to us, there had been an earlier mathematics education dissertation study of mathematical word problems as a genre (Gerofsky, 1999).

The linguistic conventions of mathematical proof can be considered a part of the genre of proof. In one study, it was found that university "students were either unaware of these conventions or unaware that these conventions applied to proof writing" and that "students did not fully understand the nuances involved in how mathematicians introduce objects in proofs" (Lew & Mejia-Ramos, 2019, p. 121). For example, mathematicians objected to "None of the sets are $\emptyset$", but students did not.

In their theoretical study, Dawkins and Weber (2017) considered values and norms of written proofs, many of which can be considered as a part of the genre of proof. Examples of norms were: (1) Justification in a proof should be deductive and not admit rebuttals. (2) Mathematical knowledge and justification should be independent of (nonmathematical) contexts, including time and author. (3) Proof is an autonomous object and not a description of the actual problem solving involved.

4.3.6 Questioning whether what mathematicians do should influence educational practices

Presumably, the main reason for investigating mathematicians' practices with respect to proof and proving is to inform the design of instruction. One thoughtful paper has analyzed whether this is wise. Using examples from the literature, the authors found "every research methodology for investigating mathematical practice is fundamentally limited and we require triangulation from multiple methods and theoretical lenses to fully understand mathematical practice." And "we highlight reasons for why mathematical practice sometimes should not inform mathematics instruction." (Weber, Dawkins, & Mejia-Ramos, 2020. p. 1063).

In general, trying to infer how to help novices become experts, from the study of experts' behavior, is fraught with difficulties. One can investigate novices' behaviors and difficulties – that's the beginning point. One can also investigate experts' behaviors--that's the theoretical endpoint, but how to get from one to the other is not easy. The general expert-novice literature often boils the problem down to observing that "It takes ten years, or 10,000 hours, to make an expert [in almost anything]." (Kendra, 2022). Mathematics education researchers are beginning to devise ways--strategies and tools--for helping university mathematics students with proving and proof comprehension, thereby making the acquisition of these important skills somewhat easier; but it seems that such learning also requires motivation and persistence. As Euclid reportedly replied to King Ptolemy, "There is no royal road to geometry."

4.4 Attempts to Help University Mathematics Students Learn How to Read and Write Proofs – Some Successes, Some Failures: Revising Proof Attempts, E-Proofs, Self-Explanation Training, Proof Frameworks

Mathematics education researchers have considered university students' difficulties with proofs and proving and how to help them learn to comprehend and construct proofs. They have considered several approaches and investigated their usefulness.

4.4.1 Revising/Rewriting proofs

Mathematics instructors commonly write comments on university students' submitted proofs, but what do students do with that feedback? One exploratory interview study (Moore, Byrne, Hanusch, & Fukawa-Connelly, 2016) asked eight undergraduate students, who had taken at least two proof-based university mathematics courses, to describe and interpret a professor's feedback on scored proofs taken from a previous study by Moore (2016). The authors found that the students tended to interpret, or misinterpret, the professor's feedback in the following ways:

(1) When the professor wrote comments on a proof, the participants correctly identified the changes, and generally, but not always, could provide some rationale for those changes.
(2) When the professor's comment was a clarification request and did not provide new text, the participants struggled to provide a rationale for the change.
(3) When they revised the proofs, the participants mostly successfully implemented the suggested changes, even when they did not fully understand the rationale for the changes. (This is a summary of the authors' findings, Moore, Byrne, Hanusch, & Fukawa-Connelly, 2016, pp. 321-322).

One can ask whether undergraduate students would have more success interpreting feedback if they were asked to revise their proofs and resubmit them for additional credit. While I did not find research on this, my experience working closely with an undergraduate real analysis instructor was that students' subsequent rewritten proofs were always deemed improved by the instructor, whether or not the students correctly interpreted the feedback.

4.4.2 Students' difficulty translating informal ideas into acceptable proofs

Students sometimes have informal ideas and arguments, yet have difficulty transforming these arguments into proofs acceptable to the mathematical community. As pointed out by Mamona-Downs and Downs (2009), "A particularly frustrating circumstance for a student is when he/she can 'see' a reason why a mathematical proposition is true, but lacks the means to express it as an explicit argument in one form or another." In their CERME-6 Working Group paper, they give three examples illustrating how students might accomplish this.

In another study, the researchers explored what it means for undergraduate to base a proof on an informal argument. They found that students pay "attention to only a small cross section of what can be carried forward from an informal argument toward the goal of creating a proof" and that "can account for some student difficulties with, and avoidance of, using their own informal arguments as a basis for proof construction." (Zazkis & Villanueva, 2016, p. 318).

4.4.3 e-proofs: *A valiant attempt to get students to understand the structure of proofs*

Even though mathematics instructors explain proofs in advanced mathematics lecture classes, students do not always know what to pay attention to or what to write in their notes (for more information, see Fukawa-Connelly, Weber, & Mejia-Ramos, 2017). Thus, most of the work of proof comprehension is done outside of class, and students' own notes are often not very helpful. This may be because, often to understand a proof and infer a warrant, one may need to consider the logical relation between two distinct lines in different parts of a proof. To help with this, a potential technological solution was devised: *e-Proofs*. These were designed to make the structure and reasoning in a proof more explicit by "graying out" less relevant parts of the proof, in order to make the structure and reasoning more explicit (Alcock & Wilkinson, 2011). Unfortunately, when tested, it was revealed "that students who studied an e-Proof did not learn more than students who had simply studied a printed proof and in fact retained their knowledge less well." This led to the conclusion that "e-Proofs made learning feel easier, but as a consequence resulted in shallower engagement and therefore poorer learning." (Alcock, Hodds, Roy, & Inglis, 2025, p. 742). The authors go on to caution that

> good pedagogical intentions do not always translate into effective interventions. It does not mean that resources like e-Proofs are never valuable—it could be, for example, that they are not good for first-time learners but are valuable resources for students who have already studied a proof independently and would benefit from clarification on aspects that they have found confusing or difficult. (Alcock, Hodds, Roy, & Inglis, 2025, p. 744).

4.4.4 Self-explanation training on how to read proofs

After considering the implications of their eye-tracking study of experts' versus novices' reading of proofs and noting how experts moved back and forth when reading proofs for understanding (see Section 4.3.1 and Inglis & Alcock, 2012), Alcock and collaborating researchers at

Loughborough University decided to try *self-explanation training* applied to mathematical proof comprehension. This is a technique that had already been tried previously with Newtonian mechanics. The materials they developed, which consisted of slides and a booklet, instructed students to identify key ideas in each line of a proof and to explain each line in terms of other ideas in the proof and their own knowledge. It was found that, in contrast to a control group who studied history of mathematics materials, students who engaged with the self-explanation training materials were better at inferring warrants, at relating different lines of a proof, and were more goal driven (Alcock, Hodds, Roy, & Inglis, 2015).

4.4.5 Proof frameworks

One can think of proof frameworks as an attempt to help beginning university mathematics students with producing proofs. We first discussed the notion of proof frameworks in our "unpacking paper". At the time, we wrote

> By a *proof framework* we mean a representation of the "top-level" logical structure of a proof, which does not depend on detailed knowledge of the relevant mathematical concepts, but which is rich enough to allow the reconstruction of the statement being proved or one equivalent to it. A written representation of a proof framework might be a sequence of statements, interspersed with blank spaces, with the potential for being expanded into a proof by additional argument. (Selden & Selden, 1995, p. 129).

Since then, we have written more extensively on proof frameworks (Selden, Selden, & Benkhalti, 2017), and others have researched similar tools such as *proof templates* (Klanderman & Satyam, 2022). However, more research needs to be done on the effectiveness of these tools.

5. Summing up and Looking to the Future[iv]

From the above, one can see that there is a rich mathematics education research literature on mathematicians' and university students' proof validation, proof construction[v], proof comprehension, proof evaluation, use of examples, and so forth, that continues to expand in a variety of new directions. I feel that part of the reason for this explosion of research has been the increasing development of new PhD programs focusing on RUME in US university mathematics departments (e.g., San Diego State University/ University of California at San Diego, established in 1993; Arizona State University established in 1997; Portland State University established in 1997; Texas State University established in 2007). Another reason for this explosion may be the introduction of new research methodologies beyond the more traditional observation, interview, and questionnaire studies to include comparative judgment studies (e.g., Davies, et al., 2020); eye-tracking studies (e.g., Inglis & Alcock, 2012); and internet surveys, particularly of mathematicians (e.g., Inglis, et al., 2013). An additional reason for this explosion of research on proof at the undergraduate level may be the creation in 2015 in Europe of INDRUM (International Network for Didactic Research on University Mathematics)[vi], which hosts biennial conferences on research in the didactics of mathematics at tertiary level and has a Working Group on Logic Reasoning and Proof.

Some promising areas for future research deal with proof assistants, and other technology, in proof and its teaching (e.g., Hanna, Reid, & de Villiers, 2019); and in research on embodied cognition, especially gestures, in proving (e.g., Kokushkin, 2022). Proof assistants can alleviate the burden of checking one's logic. Gestures can carry some of the burden of individual working memory. In addition, "Learners use collaborative gestures to extend mathematical ideas over multiple bodies as they explore, refine, and extend each other's mathematical reasoning." (Walkington, et al., 2019).

Furthermore, Alibali, Nathan, and colleagues have some intriguing results on simple proofs, done with psychology undergraduates, that indicate that gestures and speech complement one another in certain proving tasks; and that one can even predict from gestures, if they are in synch with utterances, the validity of a proof attempt. (e.g., Pier, et al., 2014). More indications for the value of studying gestures during proving comes from the research of Hortensia Soto and Michael Oehrtman (2022), who have examined the role of gestures in complex variables learning and proving. In particular, they have considered the gesture, known as an *amplitwist* for complex multiplication as a rotation and dilation, in students' and mathematicians' understanding of complex differentiation and contour integration. No doubt, other very interesting, as yet unforeseen, directions in research on undergraduates' and mathematicians' proof and proving will develop as time marches on.

References

Alibert, D. (1988). Towards new customs in the classroom. *For the Learning of Mathematics*, *8*(2), 31-35 & 43.

Alibert, D., & Thomas, M. (1991). Research on mathematical proof. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 215-243). Kluwer.

Alcock, L. (2010). Mathematicians' perspectives on the teaching and learning of proof. In F. Hitt, D. Holton, & Thompson, P. (Eds). *Research in Collegiate Mathematics Education, VII*. (pp. 63-92). CBMS Issues in Mathematics Education, Vol. 16. American Mathematical Society.

Alcock, L., Hodds, M., Roy, S., & Inglis, M. (2015). Investigating and improving undergraduate proof comprehension. *Notices of the AMS*, *62*(7), 742-752.

Alcock, L., & Weber, K. (2005). Proof validation in real analysis: Inferring and checking warrants. *Journal of Mathematical Behavior*, *24*, 125-134.

Alcock, L., & Wilkinson, N. (2011). e-Proofs: Design of a resource to support proof comprehension in mathematics. *Educational Designer*, *1*(4). http://www.educationaldesigner.org/ed/volume1/issue4/article14/index.htm

Antonini, S., & Mariotti, M. A. (2006). Reasoning in an absurd world: Difficulties with proof by contradiction. In J. Novotná, H. Moraová, M.Krátká, & N.Stehlíková (Eds.), *Proceedings of the 30th PME Conference* (Vol. 2, pp. 65–72). Prague, Czech Republic.

Bakker, A., Cai, J. & Zenger, L. (2021). Future themes of mathematics education research: An international survey before and during the pandemic. *Educational Studies in Mathematics, 107*, 1–24. https://doi.org/10.1007/s10649-021-10049-w

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

Beitlich, J. T., Obersteiner, A., Moll, G., Mora Ruano, J. G., Pan, J., Reinhold, S., & Reiss, K. (2014). The role of pictures in reading mathematical proofs: An eye movement study. In P. Liljedahl, C. Nicol, S. Oesterle, & D. Allan, D. (Eds.), *Proceedings of the 38th Conference of the International Group for the Psychology of Mathematics Education and the 36th Conference of the North American Chapter of the Psychology of Mathematics Education* (Vol. 2, pp. 121-128).Vancouver, Canada, PME. DOI: 10.13140/RG.2.1.2235.4083

Bell, A. (1976). A study of pupils' proof-explanations in mathematical situations. *Educational Studies in Mathematics*, *7*, 23-40.

Boero, P., Garuti, R., & Mariotti, M. A. (1996). Some dynamic mental processes underlying producing and proving conjectures. *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education PME-XX* (Vol. 2, pp. 121–128). Valencia.

Bowers, D. M., & Küchle, V. A. B. (2020). Mathematical proof and genre theory. *The Mathematical Intelligencer*, *42*, 48–55.

Bressoud, D. (2019). Recent research into IBL in mathematics. *Launchings, Teaching & Learning*. Mathematical Association of America. Available online at https://www.mathvalues.org/masterblog/ibl

Coppin, C. A., Mahavier, W. T., May E. L., & Parker, G. E. (2009). *The Moore Method: A pathway to learner-centered instruction*. MAA Notes Series. Mathematical Association of America.

Davies, B., Alcock, L., & Jones, I. (2020). Comparative judgement, proof summaries and proof comprehension. *Educational Studies in Mathematics*, *105*, 181–197.

Davis, P. J., & Hersh, R. (1981). *The mathematical experience*. Viking Penguin.

Dawkins, P. C., & Norton, A. (2022). Identifying mental actions necessary for abstracting the logic of conditionals. *The Journal of Mathematical Behavior, 66*, 100954.

Dawkins, P. C., & Roh, K. H. (2022). The role of unitizing predicates in the construction of logic. *Proceedings of the Twelfth Congress of the European Society for Research in Mathematics Education (CERME12)*. Bozen-Bolzano, Italy. hal-03746875v2

Dawkins, P. C., & Weber, K. (2017). Values and norms of proof for mathematicians and students. *Educational Studies in Mathematics*, *95*(2), 123-142.

DeBellis, V. A., & Goldin, G. A. (2006). Affect and meta-affect in mathematical problem solving: A representational perspective. *Educational Studies in Mathematics,* 63 (2006), 131–147.

De Villiers, M. (1990). The role and function of proof. *Pythagoras, 24*, 17-24.

Dorier, J.-L. (1998). The role of formalism in the teaching of the theory of vector spaces. *Linear Algebra and its Applications* (Vol. 275-276), 141-160. DOI: 10.1016/S0024-3795(97)10061-1.

Dubinsky, E. (1987). On teaching mathematical induction, I. *Journal of Mathematical Behavior*, *6*(1), 305-317

Dubinsky, E. (1989). On teaching mathematical induction, II. *Journal of Mathematical Behavior, 8*, 285-304.

Dubinsky, E., Elterman, F., & Gong, C. (1988). The student's construction of quantification. *For the Learning of Mathematics, 8*(2), 44-51.

Dubinksy, E., & Harel, G. (Eds.) (1992). *The concept of function*: *Aspects of epistemology and pedagogy*. MAA Notes Volume 25. Mathematical Association of America.

Durand-Guerrier, V., Boero, P., Douek, N., Epp, S., & Tanguay, D. (2012). Examining the role of logic in teaching proof. In Hanna, G. & de Villiers, M. (Eds.), *ICMI Study 19 book: Proof and proving in mathematics education* (pp. 369–389). Springer.

Durand-Guerrier, V., Hochmuth, R., Nardi, E., & Winsløw, C. (Eds.) (2021). *Research and development in university mathematics education: Overview produced by the International Network for Didactic Research in University*. Routledge.

Ellis, A. B., Lockwood, E., Dogan, F. M., Williams, C. C., & Knuth E. (2013). Choosing and using examples: How example activity can support proof insight. In A. M. Lindmeier and A. Heinze (Eds.), *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education,* (Vol. 2, pp. 265-272). Kiel, Germany: PME.

Erikson, S. A., & Lockwood, E. (2021). Investigating undergraduate students' proof schemes and perspectives about combinatorial proof. *Journal of Mathematical Behavior*, *62*, 100868.

Fischbein, E. (1982). Intuition and proof. *For the Learning of Mathematics, 3*(2), 9-18, 24.

Freudenthal, H. (1973). *Mathematics as an educational task.* D. Reidel.

Fukawa-Connelly, T. (2012). A case study of one instructor's lecture-based teaching of proof in abstract algebra: Making sense of her pedagogical moves. *Educational Studies in Mathematics, 81*, 325-345.

Fukawa-Connelly, T. (2014). Using Toulmin analysis to analyse an instructor's proof presentation in abstract algebra. *International Journal of Mathematical Education*, *45*(1), 75-88. DOI: 10.1080/0020739X.2013.790509

Fukawa-Connelly, T., Weber, K., & Mejia-Ramos, J. P. (2017). Informal content and student note-taking in advanced mathematics classes. *Journal for Research in Mathematics Education*, *48*(5), 567-579. DOI: https://doi.org/10.5951/jresematheduc.48.5.0567

Fuller, E., Weber, K., Mejía Ramos, J. P., Samkoff, A., & Rhoads, K. (2014). Comprehending structured proofs. *International Journal for Studies in Mathematics Education*, *7*(1), 1-32.

Gerofsky, S. A. (1999*). The word problem as genre in mathematics education*. [Doctoral dissertation, Simon Fraser University]. National Library of Canada. www.collectionscanada.gc.ca/obj/s4/f2/dsk1/tape7/PQDD_0027/NQ51864.pdf

Haberle, Z., Laursen, S. L., and Hayward, C. N. (2018). What's in a name? Framing struggles of a mathematics education reform community. *International Journal for Research in Undergraduate Mathematics Education.* 4:415–441. https://doi.org/10.1007/s40753-018-0079-4

Hanna, G. (1989). Proofs that prove and proofs that explain. In G. Vergnaud, J. Rogalski, & M. Artigue (Eds.), *Proceedings of the International Group for the Psychology of Mathematics Education*, Vol. II (pp. 45-51). Paris: PME Books.

Hanna, G. (1990). Some pedagogical aspects of proof. *Interchange*, *21*(1), 6-13.

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

Hanna, G. & Jahnke N. (1996). Proof and proving. In Bishop, A., Clements, M. A., Keitel-Kreidt, C., Kilpatrick, J., Koon Shing Leung, F. (Eds.), *Second International Handbook of Mathematics Education* (pp. 877-908). Kluwer.

Hannula, M. S., Leder, G. C., Morselli, F., Vollstedt, M., & Zhang, Q. (Eds.) (2019). *Affect and mathematics education: Fresh perspectives on motivation, engagement, and identity*. ICME-13 Monographs. Springer. https://doi.org/10.1007/978-3-030-13761-8

Harel, G. (1998). Two dual assertions: The first on learning and the second on teaching (or vice versa). *The American Mathematical Monthly*, *105*, 497-507.

Harel, G. (2002). The development of mathematical induction as a proof scheme: A model for DNR-based instruction. In S. R. Campbell & R. Zazkis (Eds.), *Learning and teaching number theory*: *Research in cognition and instruction* (pp. 185-212). Ablex.

Harel, G. (2008). What is mathematics? A pedagogical answer to a philosophical question. In B. Gold & R. A. Simons (Eds.), *Proof and other dilemmas: Mathematics and philosophy* (pp. 265-290). Mathematical Association of America.

Harel, G., Selden, A., & Selden, J. (2006). Advanced mathematical thinking: Past, present and future. In A. Gutierrez & P. Boero (Eds.), *Handbook of research on the psychology of mathematics education* (pp.147-72). Sense Publishers.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from an exploratory study. In A. H, Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in collegiate mathematics education. III* (pp.234-283). CBMS Issues in Mathematics Education, Vol. 7. American Mathematical Society.

Holton, D. (Ed.), (2003). *The teaching and learning of mathematics at university level: An ICMI study*. Kluwer.

Housman, D., & Porter, M. (2003). Proof schemes and learning strategies of above-average mathematics students. *Educational Studies in Mathematics*, *53*, 139-158.

Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, *43*(4), 358–390.

Inglis, M., Mejía Ramos, J. P., Weber, K., & Alcock, L. (2013). On mathematicians' different standards when evaluating elementary proofs. *Topics in Cognitive Science, 5*(2), 270-282. https://doi.org/10.1111/tops.12019

Johnson, E., Caughman, J. S., Fredericks, J., & Gibson, L. (2013). Implementing inquiry-oriented curriculum: From the mathematicians' perspective. *Journal of Mathematical Behavior*, *32*(4), 743-760.

Kendra, C. (March 29, 2022). How hard is it to become an expert? *Verywell Mind*. Downloaded from https://www.verywellmind.com, November 6, 2022.

Klanderman, S., & Satyam, V. R. (2022).Mathematical understanding and ownership in learning: affordances of and student views on templates for proof-writing. *International Journal of Mathematical Education in Science and Technology*. DOI: 10.1080/0020739X.2022.2139775

Ko, Y.-Y., & Knuth, E. J. (2013). Validating proofs and counterexamples across content domains: Practices of importance for mathematics majors. *Journal of Mathematical Behavior*, *32*, 20-35.

Kokushkin, V. (2022) The role of students' gestures in offloading cognitive demands on working memory in proving activities. PME-NA44 poster of PhD dissertation research at Virginia Tech.

Lai, Y., Weber, K., & Mejia-Ramos, J. P. (2012). Mathematicians' perspectives on features of a good pedagogical proof. *Cognition and Instruction*, *30*(2), 146-169.

Lakatos, I. (1976). *Proofs and refutations*. Cambridge University Press.

Leron, U. (1983). Structuring mathematical proofs. *American Mathematical Monthly*, *90*(3), 174-184.

Lew, K., Fukawa-Connelly, T., Mejia-Ramos, J. P., & Weber, K. (2016). Lectures in advanced mathematics: Why students might not understand what the mathematics professor is trying to convey. *Journal for Research in Mathematics Education*, *47*(2), 162-198.

Lew, K., & Mejia-Ramos, J. P. (2019). Linguistic conventions of mathematical proof writing at the undergraduate level: Mathematicians' and students' perspectives. *Journal for Research in Mathematics Education, 56*(2), 121-155.

Mamona-Downs, J., & Downs, M. (2009). Necessary realignments from mental argumentation to proof presentation. In V. Durand-Guerrier, S. Soury-Lavergne, & F. Arzarello, (Eds.), *Proceedings of the Sixth Congress of the European Society for Research in Mathematics Education* (pp. 2336-2345). Lyon: Service des publications.

Mason, J., & Pimm, D. (1984). Generic examples: Seeing the general in the particular. *Educational Studies in Mathematics*, *15*, 277-289.

McLeod, D. B. (1989). Beliefs, attitudes, and emotions: New views of affect in mathematics education. In D. B. McLeod & V. M. Adams (Eds.), *Affect and mathematical problem solving: A new perspective* (pp. 245-258). Springer-Verlag.

Mejia-Ramos, J. P., Evans, T., Rittberg, C., & Inglis, M. (2021). Mathematicians' assessments of the explanatory value of proofs. *Axiomathes*, *31*, 575–599.

Mejía Ramos, J. P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics, 79*(1), 3-18. https://doi.org/10.1007/s10649-011-9349-7

Mejía-Ramos, J. P. & Inglis, M. (2009). Argumentative and proving activities in mathematics education research. In Fou-Lai Lin, Feng-Jui Hsieh, Gila Hanna, & Michael de Villiers (Eds.), *Proceedings of the ICMI Study 19 Conference: Proof and Proving in Mathematics Education*, (Vol. 2., pp 88-93). National Taiwan Normal University.

Mejía Ramos, J. P., Lew, K., de la Torre, J., & Weber, K. (2017). Developing and validating proof comprehension tests in undergraduate mathematics. *Research in Mathematics Education, 19*(2), 130-146. https://doi.org/10.1080/14794802.2017.1325776

Mejia-Ramos, J. P., & Weber, K. (2020). Using task-based interviews to generate hypotheses about mathematical practice: Mathematics education research on mathematicians' use of examples in proof-related activities. *ZDM: Mathematics Education*, *52*(6), 1099-1112.

Mejia-Ramos, J. P., Weber, K., & Fuller, E. (2015). Factors influencing students' propensity for semantic and syntactic reasoning in proof writing: A single-case study. *International Journals of Research in Undergraduate Mathematics Education*, *1*(2), 187-208. https://doi.org/10.1007/s40735-015-0014-x

Melhuish, K., Fukawa-Connelly, T., Dawkins, P., Woods C., & Weber, K. (2022). Collegiate mathematics teaching in proof-based courses: What we now know and what we have yet to learn. *Journal of Mathematical Behavior*, *67*(7), 100986. DOI: 10.1016/j.jmathb.2022.100986

Miller, D., Engelke, N., & Weber, K. (2018). How mathematicians assign points to student proofs. *Journal of Mathematical Behavior*, *49*, 24-34.

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics, 27,* 249-266.

Moore, R. C. (2016). Mathematics professors' evaluation of students' proofs: A complex teaching practice. *International Journal of Research in Undergraduate Mathematics Education, 2*, 246-278.

Moore, R. C., Byrne, M., Hanusch, S. & Fukawa-Connelly, T. (2016). When we grade students' proofs, do they understand our feedback? In Weinberg, A., Moore-Russo, D., Soto, H. & Wawro, M. (Eds.), *Proceedings of the 19th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 310-324). Available online at: http://sigmaa.maa.org/rume/Site/Proceedings.html

Nardi, E. & Winsløw, C. (2018). INDRUM 2016 [Special issue editorial]. *International Journal of Research in Undergraduate. Mathematics. Education, 4*, 1–7. https://doi.org/10.1007/s40753-018-0074-9

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed? *Educational Studies in Mathematics*, *66*(1), 23-41. DOI 10.1007/s10649-006-9057-x

Pier, E., Walkington, C., Williams, C., Boncoddo, R., Alibali, M. W., Nathan, M. J., & Waala, J. (2014). Hear what they say and watch what they do: Predicting valid mathematical proofs using speech and gesture. In W. Penuel, S. A. Jurow, & K. O'Connor (Eds.), *Learning and becoming in practice: Proceedings of the Eleventh International Conference of the Learning Sciences* (Vol. 2, pp. 649-656). Boulder, CO: University of Colorado.

Rabin, J. M., & Quarfoot, D. (2021). Sources of students' difficulties with proof by contradiction. *International Journal of Research in Undergraduate Mathematics Education*, https://doi.org/10.1007/s40753-021-00152-x

Raman, M. (2003). Key ideas: What are they and how can they help us understand how people view proof? *Educational Studies in Mathematics, 52*(3), 319-325.

Rasmussen, C., & Kwon, O. N. (2007). An inquiry-oriented approach to undergraduate mathematics. *Journal of Mathematical Behavior*, *26*(3), 189-194.

Relaford-Doyle, J., & Núñez, R. (2021). Characterizing students' conceptual difficulties with mathematical induction using visual proofs. *International Journal of Research in Undergraduate Mathematics Education*, *7*, 1-20. DOI: 10.1007/s40753-020-00119-4

Rowland, T. (2002). Generic proofs in number theory. In S. R. Campbell and R. Zazkis (Eds.), *Learning and Teaching Number Theory: Research in Cognition and Instruction* (pp.157-183). Ablex.

Rozgonjuk, D., Kraav, T., Mikkor, K., Orav-Puurand, K., & Täht, K. (2020). Mathematics anxiety among STEM and social sciences students: The roles of mathematics self-efficacy, and deep and surface approach to learning. *International Journal of STEM Education*, *7*(46). https://doi.org/10.1186/s40594-020-00246-z

Satyam, V. R. (2020). Satisfying moments during the transition-to-proof: Characteristics of moments of significant positive emotion. *Journal of Mathematical Behavior*, *59*, 100784

Savic, Milos (2012). Proof and proving: Logic, impasses, and the relationship to problem solving. [Doctoral dissertation, New Mexico State University].

Selden, A., McKee, K., & Selden, J. (2008). The relation between affect and the proving process. *Topic Study Group 17: Research and development in the teaching and learning of advanced mathematical topics*. ICME-11, Mexico. Downloaded from *Researchgate*, November 9, 2022.

Selden, A., McKee, K., & Selden, J. (2010). Affect, behavioural schemas and the proving process. *International Journal of Mathematical Education in Science and Technology*, *41*(2), 199-215. DOI: 10.1080/00207390903388656

Selden, A., & Selden J. (1978). Errors students make in mathematical reasoning. *Bosphorus University Journal*, *6*, 67-87.

Selden, A., & Selden J. (1987). Errors and misconceptions in college level theorem proving. In J. D. Novak (Ed.), *Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics*, Vol. III (pp. 457-470). Also, available as Tennessee Technological University Tech Report No. 2003-3 at http://www.math.tntech.edu/techreports/techreports.html.

Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem?" *Journal for Research in Mathematics Education*, *34*(1), 4-36.

Selden, A., & Selden, J. (Guest Eds.) (2005). Advanced mathematical thinking [Special issue]. *Mathematical Thinking and Learning*, *7*(1).

Selden, A., & Selden, J. (2013). The genre of proof. Contribution to the chapter “Reflections on justification and proof”. In M. N. Fried & T. Dreyfus (Eds.), *Mathematics and Mathematics Education: Searching for Common Ground* (pp. 248-251). Springer. DOI: 10.1007/978-94-007-7473-5.

Selden, A., & Selden, J. (2015). A comparison of proof comprehension, proof construction, proof validation and proof evaluation. In R. Goller, R. Biehler, R. Hochmuth, & H.-G. Ruck (Eds.), *Didactics of mathematics in higher education as a scientific discipline* (pp. 340-346). Universitatbibliothek Kassel.

Selden, A., & Selden, J. (2017). An expanded theoretical perspective for proof construction and its teaching. In Dooley, T. & Gueudet, G. (Eds.), *Proceedings of the Tenth Congress of the*

European Society for Research in Mathematics Education. Dublin, Ireland: DCU Institute of Education and ERME.

Selden, J., Mason, A., & Selden, A. (1989). Can average calculus students solve nonroutine problems? *The Journal of Mathematical Behavior*, *8*, 45-50.

Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics, 29*, 123-151.

Selden, A., Selden, J., & Benkhalti, A. (2017) Proof frameworks—A way to get started. *PRIMUS*, *28*(1), 31-45. DOI: 10.1080/10511970.2017.1355858

Schunn, C. D., & Nelson, M. M. (2009). Expert-novice studies: An educational perspective. In E. M. Anderman & L. H. Anderman (Eds.), *Psychology of classroom learning: An encyclopedia.* Macmillan. DOI:10.5860/choice.47-0053.

Soto, H., & Oehrtman, M. (2022). Undergraduates' exploration of contour integration: What is accumulated? *Journal of Mathematical Behavior*, *66*, 1000963. https://doi.org/10.1016/j.jmathb.2022.100963

Speer, N., Smith, J., & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *Journal of Mathematical Behavior*, *29*(2), 99-114.

Tall, D. (Ed.) (1991). *Advanced mathematical thinking*. Kluwer.

Toulmin, S. (1958). *The uses of argument,* Cambridge University Press.

Varghese, T. (2011). Considerations concerning Balacheff's 1988 taxonomy of mathematical proofs. *Eurasia Journal of Mathematics, Science & Technology Education*, *7*(3), 181-192.

Walkington, C., Chelule, G., Woods, D., & Nathan, M. J. (2019). Collaborative gesture as a case of extended mathematical cognition. *Journal of Mathematical Behavior*, *55*, 100683. https://doi.org/10.1016/j.jmathb.2018.12.002

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, *48*, 101–119.

Weber, K. (2004). Traditional instruction in advanced mathematics courses: A case study of one professor's lectures and proofs in an introductory real analysis course. *Journal of Mathematical Behavior*, *23*(2), 115-133.

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education*, *39*, 431–459.

Weber, K. (2021). The role of syntactic representations in set theory. *Synthese*, *198*, 6393-6512.

Weber, K., & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics*, *56*(2/3), 209-234.

Weber, K., Dawkins, P., & Mejia-Ramos, J. P. (2020). The relationship between mathematical practice and mathematics pedagogy in mathematics education research. *ZDM: Mathematics Education*, *52*(6), 1063-1074.

Weber, K., & Mejia-Ramos, J. P. (2011). Why and how mathematicians read proofs: An exploratory study. *Educational Studies in Mathematics*, *76*, 329-344. DOI 10.1007/s10649-010-9292-z

Wilder, R. L. (1944). The nature of mathematical proof. *American Mathematical Monthly*, *51*, 309-323.

Yan, K., & Hanna, G. (2019). Identifying key ideas in proof: The case of the irrationality of $\sqrt{k}$. *Eleventh Congress of the European Society for Research in Mathematics Education*, Utrecht University. ffhal-02398535f

Zazkis, D., & Villanueva, M. (2016). Student conceptions of what it means to base a proof on an informal argument. *International Journal of Research in Undergraduate Mathematics Education*, *2*(3), 318–337. DOI: 10.1007/s40753-016-0032-3

---

[i] This section, although long, is necessarily somewhat sketchy and incomplete as the research on proof, proving, proof construction, proof comprehension, etc., has blossomed in so many directions in recent years. My apologies to those undergraduate mathematics education researchers whose work I have inadvertently overlooked.
[ii] The *Journal of Inquiry-Based Learning in Mathematics* (*JIBLM*) publishes course notes, user reviews, and modules. It consists mainly of materials submitted by mathematicians who teach using IBL to help others wanting to do the same.
[iii] The method of comparative judgements, long used in psychology, was implemented by asking mathematicians to make pairwise decisions about which of two student proof summaries was better.
[iv] In mathematics education research more generally, there have been recent surveys on future directions (e.g., Bakker, et al., 2021).
[v] We summarized our views on proof construction, and to some extent much of our other previous research, in a theoretical paper for the Working Group on Argumentation and Proof at Tenth Congress of the European Society of Research in Mathematics Education CERME-10 (Selden & Selden, 2017).

[vi] For more information, see Nardi & Winsløw (2018) or the edited book by Durand-Gurrier, et al. (2021).