

The Culture of Mathematical Judgment

Status, scrutiny, and the terms of participation in the age of AI

Doron Zeilberger once interrupted a lecture by Hyman Bass to announce that the result essentially rediscovered a classical theorem. His colleagues were impressed. Later that day he discovered that the hypotheses differed; he and Dominique Foata subsequently spent two summers establishing the connection properly. Public assurance had received its reward before the mathematical relationship had been understood.¹

The incident identifies an institutional danger. A claim to mastery can settle a social impression before earning an intellectual verdict. In refereeing, hiring, or research attention, that substitution can determine whether work is examined, developed, or remembered. Pettiness becomes consequential through the decisions it controls.

Mathematics has impersonal standards for whether a conclusion follows. Its communities distribute the time, instruction, criticism, and recognition through which people learn to establish conclusions. The objectivity of the first does not guarantee the fairness of the second. The important question is how authority over those resources remains responsive to mathematical reasons, especially when the people exercising it are capable of vanity, partiality, and self-deception.

AI is changing some of the costs of assistance and scrutiny. Institutions can use that capacity to widen competent participation or intensify a contest among people already recognized. The companion essay *Mathematical Understanding in the Age of AI* examines intellectual purposes and independence; here the concern is the authority to allocate a hearing, assess a contribution, and assign credit.²

1 A hearing before a verdict

A mathematical contribution has to receive attention before it can receive adequate assessment. This preliminary stage is demanding. A reader may need unfamiliar definitions, years of background, or a difficult translation between frameworks. Reputation can legitimately guide the first investment of time. A record of reliable work or a knowledgeable colleague's recommendation supplies information that a title and abstract do not.

The role of the signal should remain distinguishable from the role of the argument. An affiliation can help someone decide where to begin; it cannot establish an implication. An evaluator's experience can justify confidence in an initial assessment; it cannot make the assessment immune to correction. The distinction matters most when a decision about attention acquires the force of a decision about merit.

In his unfinished preface of December 1831, Galois complained that a referee's inability to understand could become, for the public, a verdict that the work meant nothing.³ His polemic

identifies a durable distinction: what an evaluator has established is narrower than everything a work may contain.

Three kinds of authority become entangled here. Epistemic authority supplies informed judgments. Administrative authority allocates scarce attention, money, and positions. Status authority confers esteem and helps determine whose next claim will receive a hearing. The same institution can exercise all three. Success in one role does not give unlimited warrant in the others.

A paper can be outside a journal's remit. A department can need a different specialty. A reader can lack time. These are genuine reasons for declining an investment. They become misleading when presented as demonstrations that the work lacks value, although that value was never assessed. Administrative finality should not impersonate epistemic finality.

The profession's stated standards already recognize much of this distinction. Merton's account of science emphasizes impersonal assessment, shared knowledge, and organized scrutiny; scientific reliability depends on institutions rather than exceptional personal virtue.⁴ The European Mathematical Society asks referees to judge intrinsic merit independently of authors' standing.⁵ These commitments provide the norm for the criticism. An institution that claims to support inquiry should keep its allocation of opportunities answerable to that purpose.

In a preregistered experiment at a finance journal, referees recruited through anonymous invitations evaluated the same manuscript under a Nobel laureate's name, an early-career researcher's name, or neither; rejection recommendations differed markedly.⁶ The question here is how such judgments and the opportunities they allocate accumulate. Merton's Matthew-effect analysis distinguishes credit, visibility, and resources.⁷ In Bol and colleagues' threshold study of Dutch early-career funding, applicants just above an award cutoff received more than twice as much subsequent funding over eight years as those just below it, excluding the initial award. Later application behavior differed too.⁸

A first award can make later achievement easier through time, advice, collaborators, and further resources. It is therefore a mistake to read the entire later difference as confirmation of an intrinsic difference present at the cutoff. Institutions help shape what they later assess.

Caroline Series recalls feeling accepted at Berkeley, isolated at Cambridge in 1977-78, and valued at Warwick; she also records later improvement.⁹ Lenore Blum describes a discouraging MIT entry encounter concerning her gender, Isadore Singer's intervention, and generally fair faculty treatment after admission.¹⁰ The contrasts show how access depends on decisions people can make differently.

These cases and studies concern mechanisms, not the prevalence of pettiness. Their importance lies in the power of evaluation: decisions can shape entry and development as well as reward completed work. That power calls for judgments whose scope is clear and whose grounds can be corrected.

2 The small vanities of powerful people

Ambition can sustain difficult work. Hardy candidly treated intellectual curiosity, professional pride, and the pursuit of reputation as coexisting motives.¹¹ The criticism begins when preserving a relative position takes priority over understanding: a correction becomes defeat, an unfamiliar person's insight becomes an intrusion, or uncertainty becomes a threat to a public identity built around mastery.

In the section on competitive mathematics in *Récoltes et Semailles*, Grothendieck examines his own possessiveness and selective generosity, including his difficulty receiving work from outside the people he already valued.¹² Love of mathematics and a flattering self-image can reinforce each other until another person's contribution exposes the conflict.

Writing about mathematical status games in 2010, Mike Shulman includes himself: "I dislike it, and I especially dislike the egotistical impulse in myself that it encourages." He also worries that secure fame can encourage pronouncements about specialties outside a person's competence.¹³ Status can constrain someone trying to establish a reputation; once established, it can protect them from constraints they should still accept.

Moral exceptionalism is a poor foundation for professional trust. Technical distinction warrants admiration for an achievement. It does not establish superior judgment in every domain or excuse damaging conduct toward others.

J. Peter May's memorial for Frank Adams combines exceptional competitiveness with kindness, friendship, and care for exposition. He recalls Adams asking for help with a student who shook when visiting him.¹⁴ Competitiveness and care can coexist. Institutions still have to respond to the consequences of conduct, even when the person responsible possesses substantial virtues.

The stakes are visible in ordinary conversation. A seminar needs decisive criticism, but also room to disclose a gap. When uncertainty is treated as inferiority, people have reasons to conceal what a discussion could repair. A question remains unasked because it would reveal insufficient familiarity. A prominent person's objection can settle the discussion without being examined. The room conveys information about rank while losing information about the mathematics.

Knowledge is distributed in ways that make this loss costly. The speaker may know the central construction; a listener may know a counterexample; a newcomer may notice an assumption that others have ceased to articulate. Gathering capable people creates an opportunity to combine these resources. Humiliation and performed mastery make some of them harder to use. The institutional cost of a status contest is the competence it prevents participants from exercising together.

Reputation can also protect a claim from checking. Voevodsky described an error in his work on presheaves with transfers that survived for years before Deligne's detailed examination exposed it. He also recounted his prolonged resistance to the implications of Simpson's counterexample

to work with Kapranov. His explanation links difficult arguments and authorial reputation to inadequate scrutiny, and helped motivate his commitment to formal verification.¹⁵

Voevodsky's self-correction shows how admiration can coexist with demanding examination. An unfamiliar author may have to supply unusually accessible detail; a celebrated author may receive trust through a name. The difference should remain proportionate to the evidence, rather than become an exemption from scrutiny.

Such exemptions need no conspiracy. A junior researcher may require a recommendation; an editor may value a prominent person's participation; a colleague may see little advantage in pursuing an awkward objection. Each can have an understandable reason to avoid friction. Together, these reasons can stabilize inadequate checking. Structures and character meet here: ordinary reluctance becomes institutional protection when decisions are concentrated and challenging them carries unequal costs.

Selective standards matter when an institution invokes rigor to justify its authority but relaxes it for favored people or arrangements. Rigor should govern the claims of evaluators as well as authors. Openness should shape entry routes as well as public declarations. Generosity should include accurate credit where credit threatens an established narrative. Conscious deception is unnecessary for these conflicts to matter; good intentions do not make their consequences disappear.

3 Depth, boundaries, and belonging

Technical depth can justify demanding prerequisites and compressed arguments. It cannot, by itself, justify an obscure route into the required competence or a judgment protected from challenge. The problem begins when difficulty of assessment becomes a claim that assessment should remain the possession of a small circle.

André Joyal's 2010 letter *Democracy and Mathematics* states the constructive alternative. He accepts specialization while criticizing artificial intellectual and professional barriers between fields, including pure and applied mathematics: "there is no need for gated-cities in mathematics." He calls for collective work making mathematics accessible and identifies the nLab as a project to extend, deepen, and diversify.¹⁶ The distinction is between acquiring knowledge and needing privileged access to those who possess it.

F. William Lawvere's argument about teaching gives this distinction a precise pedagogical form. Rigorous entry need not reproduce every historical construction used to justify a particular foundation. He challenges the assumption that one inherited foundational ladder supplies the only rigorous route, while offering conceptual alternatives and retaining genuine prerequisites.¹⁷ An educational tradition should distinguish what the learner needs for the next piece of work from what the tradition has grown accustomed to requiring first.

At Berkeley, Blum recalls a hierarchy that treated her work across logic, algebra, and computer science as less serious than pure logic.¹⁰ An inherited boundary needs an intellectual criterion relevant to the contribution being assessed.

A research introduction can specify the local threshold: one construction, a comparison map, or a small family of examples that makes the wider architecture intelligible. When local prerequisites remain unspecified, entry depends more heavily on personal guidance. Careful organization distributes some of that guidance publicly.

Emily Riehl criticizes mathematical language that calls arguments trivial or presumes everyone learned them in kindergarten. In the same 2020 interview, she describes algebraic topology as welcoming to young women in her experience and credits deliberate Women in Topology community-building efforts.¹⁸ The pairing matters. Alienating signals are practices, and her account credits identifiable people with making a particular community more welcoming.

There is an intellectual discipline behind accessible language. If something is immediate from a definition, name the definition. If a step uses standard machinery, identify the machinery and the hypotheses. If a question reveals a missing prerequisite, explain which one. A concise answer can remain exacting while giving the listener a route forward. Calling the question childish conveys less mathematical information and adds a social penalty to learning it.

In 2015, Shulman objected to treating homotopy type theory as the research program of one charismatic figure. He appreciated what Voevodsky's leadership had enabled, including the IAS year, without knowing whether alternatives would work as well. The picture of everyone following a dominant leader did not match his experience in algebraic topology, category theory, or homotopy type theory.¹⁹

Enabling leadership and intellectual obedience should therefore be distinguished. A leading mathematician can organize a school, identify fruitful questions, secure resources, and make unusual work possible. Those are contributions. They acquire an additional and unwarranted authority when the leader's interests are treated as the field's purposes or disagreement as evidence that a contribution falls outside serious mathematics. A healthy arrangement lets leadership increase others' options and lets independent competence revise its direction.

The portability of reasons supplies a practical test. A counterexample, failed estimate, missing hypothesis, comparison with prior work, or clearly stated project criterion gives another competent person something to assess. A judgment about an entire conceptual framework may require more context, but it can still identify what fails to connect, what known account it duplicates, or which promised consequence remains unsupported. The amount of explanation should fit the claim. Experience can be tacit; consequential decisions need enough explicit grounds to make challenge possible.

4 Credit as an intellectual record

The culture's memory influences its future. Credit tells readers whose work to consult, which ideas shaped an argument, and which kinds of labor are worth undertaking. When recognition compresses a collective achievement into a single victorious endpoint, it can lose resources for further thought as well as misdescribe the people involved.

Thurston's account of mathematical progress connects this issue to his own experiences. Early success in foliations could discourage others from participating; his later practice placed greater emphasis on communicating ways of thinking. He acknowledges the unusual trust and freedom his standing afforded him.²⁰ The institutional question is what happens when a culture values shared understanding in principle but makes the labor that creates it less visible than the result it supports.

Shimura's letters to Richard Taylor, reproduced in *The Map of My Life*, show what is at stake in preserving reasons. He complained that later summaries reduced the scope of his earlier work and that readers repeated those summaries without returning to the originals. A subsequent letter explained the mathematical considerations behind his modularity conjecture.²¹ His testimony supports a problem broader than whose name appears on a theorem: a familiar achievement can survive while the route to the idea becomes harder to recover.

A useful record distinguishes the conjecture, the proof, the crucial reduction, the correction, and the explanation that made a method reusable. Naming their different roles helps readers reconstruct dependencies and gives researchers reasons to share unfinished insights. A valuable reformulation should preserve the provenance of the work it reorganizes.

Recognition can be designed to support exchange. Daubechies recalls Bell Labs bringing colleagues from different departments together, hosting an informal unsolved-problem seminar through which she began working with Lagarias, and giving each collaborator more than half credit for joint papers in salary reviews.²² The arrangements made conversation easier and reduced a reason to experience collaboration as dividing a fixed reward. They supported generosity through the conditions of ordinary work.

Checking, finding references, reorganizing libraries, and explaining inherited constructions can benefit many people without producing a spectacular announcement. Where careers reward only the announcement, institutions discourage work on which their own intellectual life depends. Their professed commitment to collective knowledge should constrain what they recognize in practice.

No numerical substitute resolves the problem. The IMU-ICIAM-IMS report on citation statistics explains why journal averages and short citation windows are unreliable substitutes for assessment of individual mathematical work.²³ A measure records specified events; an account of significance explains what those events establish. Replacing a difficult judgment with a convenient number does not supply the missing explanation.

Two failures of administration can thus reinforce each other. Mechanical rules displace informed judgment; protected expertise makes its judgment unavailable for scrutiny. A better institution needs both competence and answerability.

5 What AI changes in the justification for authority

AI changes these questions through particular capacities, rather than through a general promise to dissolve hierarchy. Assistance can help someone prepare an argument, expose a dependency, or approach an unfamiliar subject. Independent checking can make parts of the evidence easier to examine. The institutional consequence depends on which cost has actually fallen and which judgments remain necessary.

First Proof's 2026 second batch establishes capability on specified research problems. Four public systems attempted ten problems with unpublished human solutions; across the systems, seven received at least one solution judged essentially flawless or needing only minor revisions by expert referees. Failed attempts included unsupported difficult steps and defective references. The experiment did not assess autonomous selection of important research directions.²⁴ Candidate arguments can now arise through additional routes, creating a larger need for responsible selection and examination.

A formal certificate supplies evidence of an encoded entailment relative to its declarations and assumptions; fidelity to the intended question, novelty, importance, and credit still require judgment.²⁵ Independent replay lets an inspectable artifact carry some of the burden otherwise assigned to familiarity with its author. Avigad's account of the formal turn describes how libraries make definitions and dependencies available for reuse and revision.²⁶ Status cannot repair or invalidate a checked derivation. It can still affect who interprets the statement and decides what to do with it.

The Equational Theories Project demonstrates an arrangement through which unfamiliar contributors can work together. More than fifty people with varied academic and professional backgrounds investigated relations among 4,694 equational laws. Public tasks, reviewed submissions, discussion, and automated replay supported the collaboration. Lean certified generating facts; external programs computed the full graph's closure. The infrastructure was chiefly formal and automated-theorem-proving work, with limited LLM assistance.²⁷

Its organizational lesson is explicit participation. A contributor can find a task, produce an inspectable result, receive review, and enter a shared project. Experts remain responsible for specifications, onboarding, and maintenance. Personal acquaintance need not bear the entire burden of establishing reliability. This creates a credible way to enlarge entry while preserving demanding standards: develop the means through which a contribution can be examined.

There is a complementary attribution example in the 2026 primitive-set preprint by Alexeev and collaborators. Model output suggested a Markov-chain method, developed through human constructions and checking into several proofs, including two previously open conjectures and the revised Banks-Martin conjecture. The authors distinguish model assistance, human

contributions, discarded suggestions, and responsibility for final proofs.²⁸ A precise account of who did what is possible even when the work crosses unfamiliar forms of collaboration.

A barrier justified by the cost of checking unfamiliar work should be reconsidered when that cost falls. Other grounds may remain: a claim may be wrong, unoriginal, poorly explained, or irrelevant to the project. Those reasons must be assessed on their own merits.

The same obligation applies when technology increases review burdens: new requirements and resources may be justified. In either direction, procedures should track the actual work required.

Cheaper candidate arguments can make interpretation and review relatively scarcer, increasing reliance on familiar names. The institutional choice is whether to intensify that contest or develop additional evaluators and public routes into serious work.

The sharpest test is reciprocal. If understanding is invoked to resist a theorem-only conception of AI achievement, explanation, correction, and maintainable tools should receive recognition in human careers too. If authors must disclose assistance and accept checking, evaluators should accept responsibility for their own judgments and tool use. If a company claims a machine achievement, its account should preserve the human mathematics, libraries, scaffolding, and revision on which the result depends. The standard applies across the institutions and technologies seeking authority.

6 Building a more answerable culture

Making participation easier and unsupported authority less effective requires coordinated choices about entry, review, credit, and resources. Four priorities connect the new technical capacities to the profession's stated purposes.

The first priority is a navigable route to assessment. A project can publish its questions, relevant prerequisites, examples of acceptable contributions, and a practical submission process. Review can distinguish a defective claim from a contribution outside the project's current priorities. A newcomer should be able to learn what is required before investing months in an invisible attempt. Clear expectations help contributors and conserve expert time.

Public availability alone is insufficient. Someone must undertake the work of interpretation, review, and integration. Shared task lists, contribution standards, introductory materials, and supported maintainers give an open project substance. An institution that invites contributions while providing no credible route to examination has moved the entrance without constructing the passage. Openness should be assessed through what participants can actually do.

The IMU's journal recommendations encourage impartial evaluation, grounds for decisions, and appeals, alongside screening and editorial judgment.²⁹ Explanation should fit the decision: identify a technical gap, provide a comparison for alleged redundancy, or state the relevant scope. The requirement is an accurate account of the judgment's basis.

An appeal is useful when it makes a substantive correction possible. Repeating the same procedure with no route to an additional competent assessment can preserve a decision without testing it. Independent readers, disclosed conflicts, and clear distinctions between technical and programmatic grounds help an institution remain corrigible. These arrangements should be scaled to the consequence of the decision and the feasible work of reviewing it.

Anonymizing review can help at a particular stage without repairing every earlier barrier. In a randomized ecology-journal trial, blinding reduced some location-related advantages but produced no detected difference in its effects between male- and female-authored papers; editors still knew identities.³⁰ The institutional lesson is to examine what a procedure changes, rather than assume that its label guarantees fairness.

The second priority is recognition of the labor that makes other contributions usable. Cohn's discussion of AI-generated mathematical technical debt describes valid but poorly explained arguments that leave substantial work for others, often with less recognition for the subsequent explanation. He also welcomes responsible new participants documenting AI-assisted work.³¹ The burden is real. The corresponding institutional responsibility is to value the people who discharge it.

Sanderson emphasizes explanation that reconstructs why a construction was introduced and how someone could have arrived at a method.³² Institutions can assess its effect: whether readers can locate the decisive idea, identify its limits, and use it elsewhere. Examples of use matter more than an automatic score.

Helping another researcher improve an argument or maintaining a shared resource should count through its intellectual effect. Recognition gives this work a viable place beyond the final announcement.

Avigad's August 2026 essay describes young mathematical tool builders whose work receives insufficient recognition within mathematics. His criticism addresses the profession's purposes and the range of contributions it acknowledges.³³ Institutions calling for better reasoning technology cannot consistently treat its creators' work as irrelevant to mathematical progress. The contradiction has practical consequences when the work needed for the future provides no viable path through the present career system.

The third priority is negotiated responsibility across new collaborations. Avigad's March 2026 account of sphere-packing formalization records a project that accepted AI contributions but worried that a company's separate completion would be presented as the whole achievement. Human scaffolding, appropriate credit, and reusable infrastructure mattered. He also reports the repair: the company cooperated on attribution and promised open work on improving the code, leaving the community cautiously optimistic.³⁴

Explicit project purposes make such negotiation possible. A collaboration can explain where a contribution falls short and agree on revisions. Correctness, credit, and maintainability need

distinct terms, so neither academic standing nor technological publicity determines completion by itself.

The EMS code assigns AI disclosure and responsibility to authors and referees, with protections for confidential material.⁵ Evaluators' use of tools belongs under scrutiny too.

The fourth priority is collective support. Gowers's account of his withdrawal from Elsevier explains why people can keep sustaining an arrangement they criticize: declining a request feels like letting down an editor or author they know. Public coordination can make a different choice easier.³⁵ Goodwill can stabilize a system through local obligations, so reform needs arrangements that let participants act together without abandoning one another.

Libraries, review systems, explanatory resources, and shared records likewise require time, maintenance, and recognized responsibility. Access to useful computing also needs provision if participation is to extend beyond those who can afford private resources. Supporting this work is a concrete way to turn increased capacity into additional competent participation. Depending indefinitely on unpaid enthusiasm for essential infrastructure creates a weak foundation for the expanded activity built upon it.

Several routes to assessment should remain available. Fields with extensive formal libraries have opportunities that fields with new definitions and little infrastructure do not yet possess. Making one technical format a universal entry requirement could reproduce exclusion through a new convention. The standard is appropriate inspectability: useful formal evidence, careful informal arguments, and explanations connecting the encoded claims to the intended mathematics.

Training additional evaluators is an educational task. Tao's account of AI-era mathematics connects understanding with responsibility, attribution, and education.³⁶ Teaching should develop the ability to formulate, challenge, revise, and use claims—the capacities that make broader participation substantive.

The Leiden Declaration, endorsed by the IMU, connects understanding and verification with attribution and responsibility, welcomes contributors from other disciplines, and calls for accessible participation and shared infrastructure.³⁷ Funding, appointments, review arrangements, and project standards give that commitment practical meaning. A declaration earns credibility through decisions made under its authority.

These proposals fit together. Entry materials create competence; review makes claims examinable; attribution makes contributions visible; recognition and resources sustain the people doing the work. Each supplies a condition the others need. A culture that adopts only the invitation can overwhelm reviewers. One that adopts only checking can neglect meaning and credit. One that praises maintenance without supporting it can lose the capacity on which the rest depends. Coherent reform treats the whole chain as part of mathematical practice.

7 Authority that enlarges the field

The strongest criticism of mathematical pettiness concerns its waste. An unsupported dismissal can close a route into a subject; protected reputation can delay correction; possessive credit can impoverish the record from which others learn. The acts may be small in individual intention and large in the opportunities they govern. Their seriousness comes from what they obstruct.

Expertise remains essential. A strong field needs people who can formulate questions, recognize significance, explain structures, and detect errors. It becomes stronger when those capacities can be acquired through more than one route and exercised without dependence on a single person's favor. Leadership then earns its authority partly by increasing others' competence, including their competence to disagree with it.

AI offers additional means for widening participation, alongside practices mathematics already possesses. The appropriate measure of mathematical authority is how well its reasons survive examination and how much further thought it makes possible for others.

Notes and sources

1. Doron Zeilberger, *Opinion* 29, 24 September 1998. Autobiographical account of the Bass interruption, differing hypotheses, and subsequent work with Dominique Foata.
2. *Mathematical Understanding in the Age of AI*, Tensorgami, 7 October 2026. Companion essay on intellectual purposes and independent inquiry; the present essay develops the institutional obligations of judgment.
3. Évariste Galois, unfinished preface written at Sainte-Pélagie, December 1831, [English text at MacTutor](#). French text in Tannery’s 1908 edition, pp. 25-27. His polemical account is attributed, not treated as an impartial finding.
4. Robert K. Merton, “The Normative Structure of Science” (1942), reprinted in *The Sociology of Science*, edited by Norman W. Storer (1973), pp. 267-278, especially pp. 270-278.
5. European Mathematical Society, *Code of Practice for Mathematical Publication*, approved 19 March 2025. Referee duties 3-6; author duties 3 and 8; editor duties 1 and 3. A professional standard, rather than a measurement of compliance.
6. Jürgen Huber et al., “Nobel and novice: Author prominence affects peer review”, *PNAS* 119 (2022): e2205779119. Primary full text. The anonymous-invitation comparison uses 101, 110, and 102 completed reports respectively; the outcome is a recommendation.
7. Robert K. Merton, “The Matthew Effect in Science”, *Science* 159 (1968): 56-63, doi:10.1126/science.159.3810.56. Separate discussions concern credit, communication, and resources.
8. Thijs Bol, Mathijs de Vaan, and Arnout van de Rijt, “The Matthew effect in science funding”, *PNAS* 115 (2018): 4887-4890. The subsequent-funding comparison excludes the original award; see the threshold design and analysis of later applications.
9. Caroline Series, interview in *EMS Magazine* (2022). Her comparison of Berkeley, Cambridge in 1977-78, and Warwick includes subsequent changes in the profession.
10. “Accidental activist: An interview with Lenore Blum, Part 1”, *Berkeley Science Review*, 21 May 2020. The entry encounter, Singer’s intervention, and later faculty treatment are distinct parts of her recollection.
11. G. H. Hardy, *A Mathematician’s Apology* (1940), §§7 and 29. Mixed motives and legitimate ambition are distinguished from misconduct.
12. Alexander Grothendieck, *Récoltes et Semailles*, Part I, §40, “La mathématique sportive” (mid-1980s). Mateo Carmona’s provisional transcription; §40 begins at PDF p. 256. Used for self-examination, without adopting separate allegations.
13. Mike Shulman, comment of 15 October 2010, in John Baez’s *Vanity and Ambition in Mathematics* (14 October). Shulman’s criticism includes himself.
14. J. Peter May, “Memorial Address for J. Frank Adams”, *The Mathematical Intelligencer* 12(1) (1990): 40-44; delivered 29 April 1989. Competitiveness, p. 40; the intimidated student and attempts to help, pp. 41-42; care for exposition, p. 41.
15. Vladimir Voevodsky, “The Origins and Motivations of Univalent Foundations”, Institute for Advanced Study (2014). First-person account of errors, delayed correction, reputation, and the development of his verification program.

16. André Joyal, “Democracy and Mathematics”, letter dated 1 October 2010, guest post published 4 October by John Baez. Final mathematical paragraphs distinguish specialization from artificial barriers and propose collective accessible knowledge.
17. F. William Lawvere, “Explicit Foundational Concepts in the Teaching of Mathematics”, English scan, pp. 621-626, especially pp. 621-622. The family archive attributes the English publication to Fabio Minazzi’s edited volume (Rome, 2007); the 2001 date concerns earlier Italian conference context.
18. Rachel Crowell, “Conducting the Mathematical Orchestra From the Middle”, interview with Emily Riehl, *Quanta*, 2 September 2020. Final question on social issues and questions on Women in Topology and discrimination. Her positive assessment is explicitly experiential.
19. Mike Shulman, comment of 26 May 2015, 6:25 pm, on *Voevodsky trending*. Appreciation of leadership accompanies his rejection of a portrayal of the field as one person’s program.
20. William P. Thurston, “On Proof and Progress in Mathematics”, *Bulletin of the AMS* 30 (1994): 161-177. Especially §§4-6, on proof, incentives, community, and his own contrasting experiences.
21. Goro Shimura, *The Map of My Life* (Springer, 2008), Appendix A3, pp. 176-180: letters to Richard Taylor, 25 November and 12 December 1994. Publisher’s PDF. His account of the reception of his work.
22. Allyn Jackson, interview with Ingrid Daubechies, *Celebratio Mathematica*, conducted in late 2018, section “Babies and Bell Labs.” The organizational practices are Daubechies’s recollections.
23. Robert Adler, John Ewing, and Peter Taylor, *Citation Statistics*, joint IMU-ICIAM-IMS committee report (2008), especially pp. 6-7 and 10-12. Journal citation measures and individual assessment are distinguished.
24. Mohammed Abouzaid, Nikhil Srivastava, Rachel Ward, and Lauren Williams, *First Proof Second Batch*, report dated 10 June 2026, submitted 16 June; §§1.1-1.3 and evaluation sections. Each submitted solution received two or three independent expert assessments. The seven-of-ten figure is across systems.
25. Lean language reference, “Validating a Lean Proof”, checked 11 October 2026. Validation concerns encoded statements and dependencies; inspect assumptions and independently replay proof artifacts.
26. Jeremy Avigad, “Mathematics and the formal turn”, *Bulletin of the AMS* 61 (2024): 225-240; preprint revised 4 November 2023. §§4-10 and 17 discuss libraries, dependencies, refactoring, collaboration, and machine learning.
27. Matthew Bolan, Joachim Breitner, Jose Brox, et al., *The Equational Theories Project: Advancing Collaborative Mathematical Research at Scale*, version 2, 16 December 2025. §§1.2-1.3, 3-4, and 11 distinguish the result, certificates, workflow, and limited LLM assistance. A collaborative precedent, rather than a measured reduction in status bias.
28. Boris Alexeev et al., *Primitive sets and von Mangoldt chains: Erdős Problem #1196 and beyond*, preprint, 1 May 2026. Introduction, §1.1, and §11 distinguish mathematical results and human/model contributions. One author is affiliated with OpenAI; the work includes a new proof of already solved Problem #164.
29. IMU Committee on Electronic Information and Communication, *Best Current Practices for Mathematics Journals*, June 2021, §§1.1-1.3. Recommendations on screening, evaluation, decision grounds, conflicts, and appeals.
30. Charles W. Fox, Jennifer Meyer, and Emilie Aimé, “Double-blind Peer Review Affects Reviewer Ratings and Editor Decisions at an Ecology Journal”, *Functional Ecology* 37(5) (2023): 1144-1157. Three-year

randomized trial: 3,739 submissions, 1,432 reviewed. §§2-4 distinguish ratings, revision decisions, imperfect blinding, and inferred gender.

31. Henry Cohn, “The technical debt of AI-generated mathematics”, 15 September 2026. Guest essay hosted by Terence Tao, on explanation, responsibility, and the work left to other researchers.

32. Grant Sanderson, “If math is more than proof, we need to better celebrate the rest of it”, 18 September 2026. Guest essay hosted by Terence Tao; an argument about values and recognition.

33. Jeremy Avigad, *What is mathematics now, and what should it be?*, 24 August 2026, especially §§3-5. The discussion of recognition for tool builders is an attributed professional diagnosis, rather than a population study of hiring.

34. Jeremy Avigad, *Mathematicians in the Age of AI*, dated 3 March 2026, submitted 4 March, §2, especially its concluding paragraph. His sphere-packing account records both disagreements over purposes and subsequent cooperation on credit and open revision.

35. Timothy Gowers, “Elsevier — my part in its downfall”, 21 January 2012. His account of refereeing obligations explains a coordination difficulty and the value of public collective action.

36. Terence Tao, *Mathematics in the age of AI*, submitted 17 August 2026, based on his 24 July ICM lecture. Especially the discussion of mathematical purposes, proxy measures, and §8 on responsibility, attribution, and education.

37. *Leiden Declaration on Artificial Intelligence and Mathematics*, 2 June 2026, doi:10.5281/zenodo.20302944. See commitments on understanding, verification, participation, attribution, and public infrastructure; endorsement establishes a stated institutional commitment.