

Should human beings be permitted to practice mathematics unaccompanied?

**Michael Harris
January 28, 2021**

Forty years ago, Paul Cohen enraged me by predicting that (at some unspecified future time) mathematicians would be replaced by computers. (Reuben Hersh, 2016)

Abstract: With increasing mainstream acceptance of automated proof verification in recent years, the question arises whether machine verification will soon be seen not only as convenient or useful but even obligatory, as some colleagues are already cheerfully predicting.

Moreover, more than one prominent mathematician is on record predicting that machines will replace human mathematicians in the creation as well as the verification of mathematical proofs, possibly long before the end of the present century.

This talk will explore the presuppositions implicit in such predictions and their implications for understanding the values of mathematics (as practiced by humans or not). Possible implications for the metaphysics of mathematics will also be considered (in passing).

Should human beings be permitted to practice mathematics unaccompanied?

In this talk I will try not to answer the question. I will rather explain what it means to think it is an appropriate question, and I will try to analyze what it entails. The alternative to "unaccompanied mathematics" is mathematics with a mechanical seal of approval.

If this becomes the norm, will it represent a change in the nature of mathematical practices?

Will it be more radical than the successive stages in professionalization of mathematics since the 19th century?

Will it entail sacrificing some of the core values of mathematics as currently practiced?

=(I will also try not to answer these questions, but rather to formulate them in a meaningful way. The talk is not meant to address the metaphysics of mathematical practice but the indisputable *fact* of mathematics as a way of being human.)

Warning: this is not a philosophical talk

I have no license to practice philosophy, accompanied or otherwise. (Insofar as philosophy attempts to answer questions rather than formulate new ones, I don't think its record of success is very reassuring.)

Instead, I will address the question as a mathematician, reflecting on the values that appear to be intrinsic to the practice of mathematics but that may just be a matter of "habit," following Pascal (the best account, I think, of mathematical *platonism*):

Il faut acquérir une créance plus facile qui est celle de l'habitude qui, sans violence, sans art, sans argument, nous fait croire les choses et incline toutes nos puissances à cette croyance, en sorte que notre âme y tombe naturellement. Quand on ne croit que par la force de la conviction et que l'automate est incliné à croire le contraire, ce n'est pas assez.

So philosophy should go more smoothly when the mathematician is 100% automaton!

A more up to date alternative to "habit" would be to consider mathematics a form of *larping* (live-action role-playing).

Does mathematics need a philosophy? (title by Gowers)

Matching philosophy to needs:

| Philosophical questions | Needs |
|---|---|
| What is a <i>new</i> mathematical idea? | Resist concentration of wealth and power. |
| What are the respective roles of <i>proof</i> and other factors in guiding mathematical practice? | Expand (or at least preserve) the scope of free human experience. |
| What is a <i>deep</i> mathematical idea? | Resist elitism in education. |

(If mathematics is to a subject of cultural criticism, does it belong to *high* culture?)

Does mathematics need a philosophy? (title by Gowers)

Matching philosophy to needs:

| Philosophical questions | Needs |
|---|---|
| What is a <i>new</i> mathematical idea? | Resist concentration of wealth and power. |
| What are the respective roles of <i>proof</i> and other factors in guiding mathematical practice? | Expand (or at least preserve) the scope of free human experience. |
| What is a <i>deep</i> mathematical idea? | Resist elitism in education. |

The first and last questions on the left are dismissed by most contemporary philosophical traditions, but since they dominate value judgment within contemporary mathematics they are the ones that interest me. The question in the middle is at least related to those considered by philosophy of mathematics that obscure the really important ones.

The needs contradict one another as well as the values expressed on the right...

Is policing the (elite) boundaries of mathematics (gatekeeping) part of mathematical practice?

Background on formal verification

Automated proof verification goes back to the early history of computer science.

It should be possible almost to eliminate debugging. . . . Instead of debugging a program, one should prove that it meets its specifications, and this proof should be checked by a computer program.

John McCarthy (who may have coined the term *artificial intelligence*), 1962.

There was an active community on the border between computer science and mathematics already in the 1990s. But I choose to start its history as an acceptable option within mainstream mathematics with Tom Hales's *Flyspeck project*, designed to provide a formal verification of his proof of the Kepler Conjecture, which involved too much software to be amenable to human verification. Subsequent successes of formalization were directed at long human proofs (the odd-order theorem in group theory) as well as hybrid proofs (the four color theorem).

Mathematical **practices** of the future

A more recent step toward mainstream acceptance is the [Xena project](#), led by Kevin Buzzard at Imperial College London.

The goal of the Xena Project is to get mathematicians, especially undergraduates, using computer proof verification software. This software can be used to check proofs of undergraduate level theorems, solve problem sheet questions and so on. The software turns maths questions into levels of a computer game. The main software we use is called [Lean](#), and we mathematicians cannot get anywhere without also using [mathlib](#), Lean's maths library.

Buzzard's talk on YouTube is entitled [The Future of Mathematics?](#)

I now clearly understand that software such as Lean is part of the inevitable future of mathematics... It's only a matter of time before mathematicians start waking up...

Will philosophers of mathematical practices wake up when the practices on which they base their philosophy become obsolete?

Libraries

An immediate goal of Xena is to formalize proofs of theorems that can then be collected in a library. Being formalized they are stamped as reliable and can then be quoted, sparing the need to formalize from scratch. Perhaps the library will also be searchable, like a database, or the internet as a whole (as I am doing to prepare this talk).

Even better: the library can be asked (like Siri): is there already a theorem that looks like [fill in the lemma that you need]? This is a major activity within mathematical practice, except that one asks "experts" if one hasn't managed to find the answer in the published record.

Is this significantly different from the Stacks project, an important online library of (rigorous but not formal) proofs of basic facts in algebraic geometry? One difference is that the Xena library should be interactive... "You can't Google," as Buzzard says.

Proof assistants (also called interactive theorem provers) are increasingly used in academia and industry to verify the correctness of hardware, software, and protocols. However, despite the trustworthiness guarantees they offer, most mathematicians find them too laborious to use.

*The ultimate aim is to develop **a proof assistant that actually helps mathematicians**, by making them more productive and more confident in their results.*

Too many proofs look more like LSD trips than coherent mathematical arguments

(from *Lean Forward*, PI Jasmin Blanchette, Amsterdam)

This was written, however, by a computer scientist, and computer scientists are notoriously enthusiastic about proof assistants.

Buzzard points out that Lean will reject the intuitive proof that the derivative of sine is cosine, for example, and that will help undergraduates understand what a proof is. I guess (but I haven't checked) that for most of the 18th century the argument using triangles *was* the proof. If we see the 19th century proof as "better" it is incumbent upon us to explain why.

Leibnizian proof

For mathematical formalists, a **proof** is a sequence of valid formulas in a symbolic language, each obtained from the previous one (more likely several) by a legal transformation (rule of deduction) or the statement of an axiom.

A **theorem** is then any valid formula that appears as the final statement in a sequence.

Here is a recent (militant but poorly substantiated) attempt to ventriloquize mathematics in support of this position:

an informal mathematical statement is a theorem if and only if its formal counterpart has a formal derivation. Whether or not a mathematician reading a proof would characterize the state of affairs in these terms, a judgement as to correctness is tantamount to a judgment as to the existence of a formal derivation, and whatever psychological processes the mathematician brings to bear, they are reliable insofar as they track the correspondence.

(Avigad, 2020)

How this could be a mathematical position

Avigad mainly quotes philosophers in support of his contention that this is the "standard view" of how mathematics really works. In fact, it amounts to the position that this is how mathematics **should** be understood.

It's not completely hopeless. On this view, the set of all proofs is a subset of the set of all sequences of symbolic formulas, and the set of all theorems has a similar description. In this way, proofs and theorems can be studied as mathematical objects, like simple groups, or 2-dimensional manifolds, or elliptic curves with rational coefficients, all of which have been *classified* (classification is given as a theorem).

The classification of theorems is undecidable. But that's not why this picture of proofs and theorems is not current among mathematicians, even (no matter what philosophers like Avigad contend) as an ideal. Nor is it merely because this view is the natural starting point for a deduction that leads to the conclusion that human mathematicians are **dispensable** for mathematics.

(And if this is the only epistemologically defensible position — as well may be the case — maybe epistemology is a dead end for mathematics.)

The real problem is that mathematics is concerned (and not only in hiring committees) with *normative judgments*, and the ideal proof is one that attracts positive judgments.

These can be:

- Ineffable (aesthetic); for example, Poincaré's *je ne sais quoi*. (Elitist!)
- A guide to *understanding* (also ineffable); for example, Atiyah wrote "*the proof is the check that I have understood...it isn't the primary thing at all.*" (Self-absorbed!)
- Influential (the language of prize nominations and Wikipedia pages); for example "opened new perspectives..." (Self-perpetuating! and possibly Eurocentric!)
- Also, one hopes that the proof is *valid* in some approximation to the Leibnizian sense.

Philosophy worries about the last of the list, which is almost always a routine matter for mathematicians, but has nothing much to say about the first three, for which mathematicians are really in need of help, and tends to dismiss them as irrelevant.

Automated proof checkers are also unlikely to help with the first three.

A different view of proof

To quote myself:

Unlike those of my colleagues who predict that computers will soon replace human mathematicians by virtue of their superior skill and reliability in proving theorems, I am inclined to respond that the goal of mathematics is to convert rigorous proofs to heuristics. The latter are in turn used to produce new rigorous proofs, a necessary input (but not the only one) for new heuristics.

*If asked to reproduce the proof [of FLT] as a sequence of logical deductions, they would undoubtedly have come up with 10 different versions.... Nevertheless — and this is what is missing from the logicians' account of proof — each of the 10 would readily refer to his or her proof as Wiles's proof. They would refer in the same way to the proofs they studied in the expository articles or in the graduate courses they taught or attended. And though each of the 10 would have left out some details, they would all be right.... **a proof like the one Wiles published is not meant to be treated as a self-contained artifact.** On the contrary, Wiles's proof is the point of departure for an open-ended dialogue that is too elusive and alive to be limited by foundational constraints that are alien to the subject matter. (emphasis added)*

In other words: even a philosopher shouldn't get too hung up on individual proofs. Bear this in mind when reading this quote from Buzzard's 2019 talk to Microsoft:

Beating humans

- Clear: tools such as Lean will one day help us mathematicians search for theorems in the literature, and help us to prove theorems. These tools may also change the way we teach.
- Possible: tools such as Lean will begin to do research semi-autonomously, perhaps uncover problems in the literature. Maybe these tools will replace research mathematicians.
- In April, Christian Szegedy from Google told me that he believes that computers will be beating humans at math within ten years.

I specifically want to understand what it can possibly mean to "beat humans" and how many ostensibly human activities are susceptible to such "beating." AI is already supposed to "beat humans" at driving and identifying tumors, some industrial processes, and games like chess and go. Also software verification.

What would it mean for AI to "beat humans" at *convincing and explaining*? At *social organization* or *linguistic communication* or *general "conversation"*? How about *philosophy*? Are these inseparable from mathematical practice?

How much of this can be formalized?

And how much can be accomplished without humans?

*While any list of the functions of proofs in mathematics is bound to be incomplete, we discern a number of regularities in the lists available in the literature. For Camilla Gilmore and colleagues (2018, section 9.1), the following main roles for proofs are identified: to verify the truth of a statement/theorem; to explain why the theorem is true; to introduce new ideas and techniques; and to systematize results into a coherent theory of definitions and other results. Paul Auslander (2008) highlights the following functions/roles: proof as certification within the relevant mathematical community; proof as explanation; proof as exploration leading to new insights, techniques or ideas; and proofs as justification of definitions. Michael De Villiers (1999) discerns verification, explanation, systematization, discovery, intellectual challenge, and communication as the roles of proofs. Reuben Hersh (1993) emphasizes **convincing and explaining** as the main function of proofs, as does John Dawson (2006), for whom a proof's "purpose is to convince those who endeavor to follow it that a certain mathematical statement is true (and, ideally, to explain why it is true)" (p. 270). J. A. Robinson also emphasizes the roles **of convincing and explaining** (as discussed in Detlefsen, 2008). On Paul Ernest's (1994) dialogical conception, proofs are viewed as a critical component of the **general 'conversation'** that is mathematics, and thus essentially as communicative, discursive entities. Yehuda Rav (1999) stresses the creative dimension of proofs in giving rise to new methods, tools, strategies and concepts; the same holds for Lakatos (Tanswell, 2018). We thus obtain the following list of (partially interrelated) roles for proofs in mathematical practices...*

C. Dutilh Novaes, *The Dialogical Roots of Deduction*

Replacing humans

It's conceivable that someday, computers will replace humans at all aspects of mathematical research — but it's also conceivable that, by the time they can do that, they'll be able to replace humans at music and science journalism and everything else!

(Scott Aaronson, in *Scientific American*, 2020)

Again, what is this talk of "replacing humans"? Assuming humans are indispensable, it can't just be for industrial production, driving vehicles (with what inside?), performing music (for what audience?) or winning games, in any of which we can imagine machines "beating" humans. So the value of mathematics has to be sought elsewhere — as a *way of being human*. In finding the words to explain this we need to avoid invoking appeals to elite values ("art for art's sake") while remaining faithful to the values of the practice.

I envy the humanists, like Gayatri Spivak, who can defend their practice with sentences like this (from the course we taught together last fall):

The digital will never capture the contingent.

Are mathematicians attached to the contingent? For purely nostalgic reasons?

The argument from intrinsic interestingness moreover has an air of trivial hobbyism about it: it makes humanistic inquiry out to be an activity suitable for someone who has no deep commitments or beliefs, and simply needs to be distracted by something or other in order not to grow completely dissolute. If this were all humanistic inquiry were, it would indeed be the special purview of an idle elite, as many already suspect it is.

Here is why I actually think humanistic inquiry should be defended: because it elevates the human spirit. Nothing is interesting or uninteresting in itself in a pre-given way. What is of interest in studying a humanistic object is not only the object, but the character of the relation that emerges between that object and oneself. What emerges from humanistic inquiry is thus best understood as an I-Thou relation, rather than an I-It relation.

Justin E. H. Smith, *What are the Humanities?*

I prefer to put "oneself" in the plural, but apart from that, can this be translated into a defense of mathematics as a human activity? Jacobi, and later Hilbert and Dieudonné, also invoked the "human spirit."¹ Has it been killed by neoliberalism? Are there only "it-it" relations?

¹Hilbert: *Die Ehre des menschlichen Geistes, so sagte der berühmte Königsberger Mathematiker JACOBI, ist der einzige Zweck aller Wissenschaft.* Thanks to Juliette Kennedy for this quotation.

People vs. things

We must rapidly begin the shift from a "thing-oriented" society to a "person-oriented" society. When machines and computers, profit motives and property rights are considered more important than people, the giant triplets of racism, materialism, and militarism are incapable of being conquered.

Martin Luther King, Jr., *Beyond Vietnam* speech, 1967

Rescuing the "human spirit" from bombast

The scientist... must appear to the systematic epistemologist as a type of unscrupulous opportunist: he appears ... as idealist insofar as he looks upon the concepts and theories as free inventions of the human spirit (not logically derivable from what is empirically given)...

Albert Einstein, 1949

Philosophers have an immense repertoire of presentable approaches to an expression that has been rejected as imprecise and outmoded but that remains as shorthand for the accepted intrinsic value of human life and activity (beyond the 29 appearances in the *Stanford Encyclopedia*, like the one just quoted, mostly referring to idealist philosophers of the 18th and 19th century).

It's questionable whether post-human philosophy is even possible; but philosophy of mathematics in particular seems to depend on some version of the rational subject, either collective or individual.

Seeing mathematics as a *way of being human* (like *larping*) is the starting point for a philosophy of mathematics rooted in ethics rather than metaphysics. Even formalism, which surrenders the epistemological questions to a mechanical process, does not view mathematics as a game that plays itself; but its ethics is underdeveloped.

One's position on the possible future mechanization of proof is a function of one's view of mathematics itself.

Is it a **means to an end** that can be achieved as well, or better, by a competent machine as by a human being? If so, **what is that end**, and why are machines seen as more reliable than humans?

Or is mathematics rather **an end in itself**, a human practice that is pursued for its intrinsic value? If so, **what could that value be**, and can it ever be shared with machines?

I believe the right question to ask, respecting all ornament, is simply this: Was it done with enjoyment—was the carver happy while he was about it? It may be the hardest work possible, and the harder because so much pleasure was taken in it; but it must have been happy too, or it will not be living. How much of the stone mason's toil this condition would exclude I hardly venture to consider, but the condition is absolute.

Ruskin, *The Seven Lamps of Architecture* (1849)

Are machine mathematics and human mathematics qualitatively different?

My impression is that computers cannot read human-generated mathematics. I think you actually have to write it properly. To get mathematicians to do that, you have to teach mathematicians to write in this programming language... there has to be a new culture.

([Buzzard](#), 2019, 54'20")

In developing theorem-proving software, practitioners endowed mathematical objects with computational properties - they were translated into dynamic, algorithmic, discrete things.

(Stephanie Dick, 2015)

Question: is that what they were already?

These tools endowed existing mathematical systems and ideas with new and different meaning.

Question: Did the meaning change?

Should human beings be **permitted** to practice mathematics **unaccompanied**?

The word in **red** is a reminder that, although the discussion of automated proof verification tends to be framed in the language of *truth*, the underlying question is one of mathematical *ethics*. This in turn can be understood as an ethics of mathematical *practice*, though it can also be construed as an ethics of mathematics *as such*, which seems to be the starting point for most mainstream (analytic, logical empiricist) philosophy of mathematics.

The word in **blue** also points to ethical concerns, but it also raises questions of authority and power. Who decides what is and is not permitted? Consider the analogous question of preparing manuscripts for publication. Before word processing there were professional technical typists. Now this job has disappeared (deskilling) and mathematicians are expected to submit manuscripts in Latex (speed-up, rationalization). The decision seems to have been made by publishers, with the acquiescence of the *leadership* of the mathematical community, with professional associations playing a mediating role. (Compare Twitter's rules.)

Should human beings be permitted to practice mathematics **unaccompanied**?

The word in blue also alludes to the prediction that preparing manuscripts in a format ready for automated proof verification will ultimately become as easy as typing in Latex. When, as suggested by the word in white, this becomes an obligation (speed-up, rationalization), the painstaking part of the refereeing process will be left to the machine (deskilling). This deskilling may seem welcome:

[Voevodsky's new system] hastens the day when our mathematical literature has been verified mechanically and referees are relieved of the tedium of checking the proofs in articles submitted for publication. (Dan Grayson, 2017)

Mathematicians may be concerned, however, that no one other than the author and the machine (ultimately the same individual) will need to "understand" the proof. Will that matter? (Buzzard pointed out that most of us don't know how our telephone works. Is that an adequate analogy?)

Should human beings be permitted to practice the saxophone **unaccompanied**?

We're going to digitize mathematics and it's going to make it better.

(Kevin Buzzard, in *Quanta*, 2020)

The header alludes to the possibility that formalization will be treated as an *aesthetic* imperative, perhaps disguised as a methodological imperative. Bourbaki established the principle that superfluous hypotheses should be ruthlessly eliminated.

Grothendieck extended this principle in a categorical framework, so that all assertions should be visibly independent of extraneous choices (or alternatively, inevitable dependence should be part of the assertion). This is methodologically extremely powerful, one of the algebraic geometer's main tools. Making type-theoretic considerations sufficiently explicit to be legible to Lean could be the aesthetic equivalent of the Centre Pompidou's *exposed services*.

For Buzzard, however, "better" is a *moral* imperative, where the commitment to *truth* is primary.

Should human beings be **permitted** to practice radiology(surgery) **unaccompanied**?

Current promotion of AI includes the claim that machine learning is more accurate at recognizing malignant tumors than human radiologists. So practicing medicine unaccompanied by machines will eventually be considered a form of malpractice.

Failure to use AI for verification of strategic software may in the future in some circumstances lead to criminal liability.

Does any of this apply to mathematics? Being able to think mathematically is clearly indispensable to face novel challenges, like pandemics (but some of these challenges are also arguments for dispensing with humans). The argument is often made that it's important for human survival (not so much the human spirit) to keep ready for emergencies a community able to think in a mathematical way. I find this argument awkward because emergencies don't require (much) number theory and the self-image of mathematics does not resemble that of the fire brigade.

The importance of heretical thinking

$ZF \quad X=Y \Leftrightarrow \forall a((x \in X) \Leftrightarrow (x \in Y))$

$a=b$ (same type)
 $is\ a\ type$
 $out\ of\ line\ a \leftrightarrow b$
 $\alpha \doteq a=b$

$First \begin{cases} \pi_n(S^n) = \mathbb{Z} \\ H_x(X \times Y, F) = H_x(X, F) \oplus H_x(Y, F) \end{cases}$

$p_1^* \cdot p_2^* \quad \alpha \quad \beta$

$C \xrightarrow{\cong} D \quad A \quad K(A)$
 $NC \xrightarrow{h \cong} ND \quad A \quad [A]$
 $\cdot A' \rightarrow A \rightarrow A''$

Pause (k)

38:52 / 47:02

What do we mean by "equal" - Pierre Deligne

18,249 views • Sep 12, 2018

404 5 SHARE SAVE ...

https://youtu.be/WfDcrN5_1wA?t=2333 (two minutes)

Are human flaws perhaps indispensable for mathematics?

The sociologist of science Harry Collins has argued that because machines are not social beings, much of human knowledge will be inaccessible to them as knowledge is inherently social in character. According to Collins, machines will only ever be capable of participating in knowledge-making in domains where human activity in them has been mechanized and rule bound already in advance. Game playing and mathematics were such domains and that is what made them viable choices for Artificial Intelligence research.

Stephanie Dick, 2015 Harvard dissertation

So we can conclude from this brief tour that computer scientists are convinced that mathematics is mechanized and rule-bound, that some mathematicians (and some sociologists!) find this image of mathematics appealing, and that the propagation of this view is facilitated by the neoliberal compulsion to evaluate everything in terms of its role in the market and by an ideology of progress founded on technological determinism. The two factors overlap; Buzzard (I should stress he is a friend!) asked on my blog whether I might not be a *Luddite*, as if theorems were of the same nature as textiles.

Examples of "heretical" thinking (if there is time)

Buzzard claims he defined a complex number (within **Lean**, a Microsoft product) as a pair of real numbers. This definition (or a more elaborate definition) may conflict with the choice in a different proof assistant, say **Spleen** (by Google, for example)! The following comment reads simultaneously as an argument for keeping humans around and as a recipe for what Jaron Lanier calls "lock-in."

*I can't get [undergraduates] to formalize the definition of a p -adic Galois representation, because they don't know what it is. ... you shouldn't try and teach somebody some new and complicated mathematics while simultaneously getting them to formalize it. ... You should get them to formalize stuff that they understand. **First they understand it**, then they formalize it, and then they really understand it, and it stays forever. (Buzzard, around 43'30", my emphasis)*

Weil cohomology

(Why I have not been able to organize my thoughts for this talk...)

By way of summary

Mainstream philosophy of mathematics promotes the perspective that proof is the primary purpose of mathematical practice and that a mathematician's proof is an approximation to an ideal formal proof.

Philosophy of mathematics aims to provide an idealized account of mathematics, a normative practice for acquiring knowledge; the data to be explained is the body of mathematical knowledge, made manifest by the history of mathematics and its literature.

(Avigad)

Predictions that computers will "beat" (or "replace") humans at mathematics only make sense in terms of this perspective.

There is no meaningful barrier between such predictions and the predictions that computers will "replace humans" altogether (and the belief that this is in some sense desirable).

This is consistent with the Horkheimer/Adorno vision of mathematics:

enlightenment ... equates thought with mathematics. Thought is reified as an autonomous automatic process aping the machine it has itself produced so that it can finally be replaced by the machine... mathematics made thought into a thing, a tool, to use its own terms."

Not coincidentally, this perspective is compatible with the neoliberal identification of value with value for the market.

In the words of Matthew Handelman, for the Frankfurt school, in particular the authors of *Dialectic of Enlightenment*,

if philosophy became the formulae of mathematics, then, the critical theorists believed, thought would be impotent to resist and, therefore, complicit in the authoritarianism engulfing Europe.

Fortunately, "the body of mathematical knowledge, made manifest" by its practice, demonstrates the limitations of the formalist perspective. So it is possible to practice mathematics and not be complicit in authoritarianism, not to mention human extinction.