

Rescuing Philosophy

Michael Johnson

Qualia Research Institute

San Francisco, CA

MIKE@QUALIARESEARCHINSTITUTE.ORG

I.

Philosophy has lost much of its energy, focus, and glamor in our modern era. What happened?

I'd suggest that five things went wrong:

1. Historical illegibility. Historically, ‘philosophy’ is what you do when you don’t know what to do. This naturally involves a lot of error. Once you figure out a core ontology and methodology for your topic, you stop calling it ‘philosophy’ and start calling it ‘science’, ‘linguistics’, ‘modal logic’, and so on. This is a very important, generative process, but it also means that if you look back at the history of philosophy, you basically only see ideas that are, technically speaking, *wrong*. This gives philosophers trying to ‘carry on the tradition’ a skewed understanding of what philosophy *is*, and how to do it.

2. Evaporative cooling. The fact that the most successful people, ideas, ontologies/methodologies, and tools tend to leave philosophy to found their own disciplines leads to long-term problems with quality. We can think of this as an evaporative cooling [15] effect where philosophy is left with the fakest of problems, and the worst, most incoherent, and confused ways to frame what real problems are left.

3. Lack of feedback loops. Good abstract philosophy is *really hard* to do right, it’s hard to distinguish good philosophy from the bad, and the value of doing philosophy well isn’t as immediately apparent as, say, chemistry. This leads to ‘monkey politics’ playing a large role in which ideas gain traction, which in turn drives a lot of top talent away.

4. Professionalization. Turning metaphysical confusion into something clear enough to build a new science on tends to be a very illegible process, full of false-starts, recontextualizations, and unpredictable breakthroughs. This is really hard to systematically teach students how to do, and even harder to plan an academic budget around. As philosophy became regularized and professionalized—something that you can have *a career* in—it was also pushed toward top-down legibility [5]. This resulted in less of a focus on grappling with metaphysical uncertainty and more focus on institutionally-legible things such as scholarship, incremental research, teaching, and so on. Today, the discipline is often taught and organized academically as a ‘history of ideas’, based on how past philosophers carved vari-

ous problem-spaces.

5. Postmodernism. Philosophy got hit pretty hard by postmodernism [6]—and insofar as philosophy was the traditional keeper of theories of meaning, and insofar as postmodernism attacked all sources of meaning, philosophy suffered *more* than other disciplines. Likewise, academic philosophy has inherited all the problems of modern academia [7], of which there are many.

I’m painting with a broad brush here, and I should note that there are pockets of brilliant academic philosophers out there doing good, and even heroic, work *in spite* of these structural conditions. But I don’t think many of these would claim they’re happy with modern academic philosophy’s structural conditions [12] or trajectory.

And this *matters*, since *philosophy is still necessary*. There’s *a lot* of value in having a solid philosophical toolset, and having a healthy intellectual tradition of being mindful about ontological questions, epistemology, and so on. As David Pearce often points out, there’s no clean way to abstain from thorny philosophical issues: “The penalty of *not* doing philosophy isn’t to transcend it, but simply to give bad philosophical arguments a free pass.”

II.

So philosophy is broken. What do we do about it?

My friend Sebastian Marshall describes the ‘evaporative cooling’ philosophy has undergone, and suggests that we should try to *rescue* and *reclaim* philosophy [10]:

‘So, this bastardized divorced left-behind philosophy will be here to stay in some form or fashion. We can’t get rid of it. . . but it’s also not necessary to get rid of it.

Turning to better news, even in mainstream philosophy, there are still sane and sound branches doing good work, like logic (which is basically math) and philosophy of mind (which is rapidly becoming neuroscience but which hasn’t yet evaporatively cooled out of philosophy).

It wouldn’t take very many people reclaiming the word philosophy as a love of wisdom to begin to turn things around.

Genuinely good philosophy is happening all over the place—though it’s rarely people in fields that don’t fight back at all. Indeed, you see computer programming and financiers doing some of the best philosophy now—Paul Graham, Eliezer Yudkowsky, Ray Dalio, Charlie Munger, Nassim Taleb. When the computer scientist get something wrong, their code doesn’t work. When the financier gets something wrong, they lose a lot of money. Excellent philosophers still come of the military—John Boyd and Hyman Rickover to name two recent Ameri-

cans—and they come out of industrial engineering, like Eli Goldratt.

That these people are currently not classified as philosophers is simply an error—let the people doing uselessness in the towers call themselves “theoretical screwaroundists” or whatever other more palatable name they might come up with for themselves; genuine philosophy is alive and well, even as the word points to decayed and failing institutions.’

There would clearly be enormous benefits to reclaiming the word “philosophy” for serious generative work. But I worry it’s going to be really hard.

Words have a lifecycle—often, they start out full of focus, wit, and life, able to vividly convey some key relationship. As time goes on, however, they lose this special something as they become normalized and regress toward the linguistic mean. Part of being a good writer is being in-tune with what words and phrases still have life, and part of being a *great* writer (like Shakespeare [9]) is minting new ones. My sense is that “philosophy” doesn’t have much sparkle left, and it may be preferable to coin a new word.

Unfortunately, I don’t know of any better words that would encapsulate everything we’d want, and it may be very difficult to rally people behind a new term unless it’s *really* good. Even though academic philosophy is in terrible shape, the term ‘philosophy’ is still an effective Schelling point; still prime memetic real-estate. So, pending that better option, I think Sebastian’s right and we do need to do what we can to rescue and reclaim philosophy.

III.

How do we rescue philosophy?—I think we need to think about this both in terms of *individual tactics* and *collective strategy*.

Individual tactics: survival and value creation in an unfriendly environment

Essentially, those that wish to make a notable, real, and durable contribution to philosophy should understand that association with academia is a double-edged sword. On the plus side, it can give people credibility, access, and fellowship with other academics, apprenticeships with established thinkers, maybe a steady income, and a great excuse to engage deeply with philosophy. On the other hand, by going into academic philosophy someone is essentially granting an unhealthy, partially moribund system broad influence over their local incentives, memetic milieu, and aesthetic. That’s *a really big deal*.

A personal aside: I struggled with how to navigate this while writing *Principia Qualia* [8]. Clearly a new philosophical work on consciousness should engage with other work in the space—and there’s a lot of *good* philosophy of mind out there, work I could probably use and build upon. At the same time, if philosophy’s established ways of framing the problem of consciousness could lead to a solution, it would’ve been solved by now, and by using someone else’s packaged ontology, I’d be at risk of importing their confusion into my foun-

ation. With this in mind I decided that *being aware* of key landmarks in philosophy was important, but *being uncorrelated with philosophy's past framing* was equally important, so I took a minimalist first-principles approach to building my framework and was very careful about what I imported from philosophy and how I used it.

Collective strategy: Schelling points and positive-feedback loops

The machinery of modern academic philosophy is going to resist attempts at reformation, as all rudderless bureaucratic entities do, but it won't be proactively hostile about it, and in fact a lot of philosophers *desperately want* change. This means people can engage in open coordination on this problem. *I.e.*, if we can identify Schelling points and plant rallying flags which can help coordinate with potential allies, we could probably make a collective push to fix certain problems or subfields (my sources say this sort of 'benign takeover' is already in motion in certain departments of bioethics).

Ultimately, though, fixing philosophy from within probably looks like a better option than it actually is, since (1) entryism is sneaky, always has a bad faith component, and is never as simple as it sounds (if nothing else, you have to fight off other entryists!), and (2) meme flow always goes both ways, and a plan to fix philosophy's norms faster than its bad norms subvert us is inherently risky. Plenty of good people with magnificent intentions of fixing philosophy go into grad school, only to get lost in the noise, fail to catalyze a positive-feedback-loop, burn out, and give up years later. If you're going into academic philosophy anyway, then definitely try to improve it, but don't go into academic philosophy *in order* to improve it.

Instead, it may be better to build institutions that are separate from modern academic philosophy, and *compete against it*. Right now, academic philosophy looks "too big to fail"—a juggernaut that, for all its flaws, is still the go-to arbiter of success, authority, and truth in philosophy. And as long as academic philosophy can keep its people stably supplied with money and status, and people on the outside have to scramble for scraps, this isn't going to change much. But nothing is forever [11] and there are hints of a shift [14], the world *needs* better alternatives, and now is a great time to start building them.

In short, I think the best way to fix philosophy may be to to build new (or revive ancient) competing metaphors for what philosophy *should* be, to solve problems that modern philosophy *can't*, to offer a viable *refuge* for people fleeing academia's dysfunction, and to *make academia come to us* if it wants to stay relevant.

IV.

This is essentially what we're working toward at the Qualia Research Institute: building something new, *outside* of academic philosophy in order to avoid its dysfunction, but still very much *focused on* a core problem of philosophy.

I see this happening elsewhere, too: LessWrong [2] is essentially a “hard fork” of epistemology, with different problem-carvings, norms, and methods, which are collectively slowly maturing into the notion of executable philosophy [1]. Likewise, Leverage Research [3] may be crazy, but I’ve got to give them credit for being crazy in a novel and generative way, one which is uncorrelated with the more mundane, depressing ways modern academic philosophy and psychology are crazy. Honorable mentions include Exosphere, an intentional community I’m pretty sure Aristotle would have felt right at home in, and Alexandros Pagidas [13], a refugee from academic philosophy who’s trying to revive traditional Greek-style philosophical fight clubs [4] (which, to be honest, sound kind of fun).

There are a lot of these little seeds around. Not all of them will sprout into something magnificent. But I think most are worth watering.

References

- [1] Executable philosophy, . URL https://arbital.com/p/executable_philosophy/.
- [2] LessWrong, . URL <https://www.lesswrong.com/>.
- [3] Leverage Research, . URL <https://www.leverageresearch.org>.
- [4] The Philosophical Fight Club, . URL <http://thephilosophicalfight.club/>.
- [5] S. Alexander. Book Review: Seeing Like A State, Mar. 2017. URL <https://slatestarcodex.com/2017/03/16/book-review-seeing-like-a-state/>.
- [6] M. Doria. The Unreasonable Destructiveness of Political Correctness in Philosophy. *Philosophies*, 2(3):17, Sept. 2017. doi: 10.3390/philosophies2030017. URL <https://www.mdpi.com/2409-9287/2/3/17>.
- [7] J. Hawks. Funding must make room for exploration, May 2017. URL <http://johnhawks.net/weblog/topics/metascience/funding/exploration-necessary-perlmutter-2016.html>.
- [8] M. Johnson. Principia Qualia, Nov. 2016. URL <https://opentheory.net/2016/11/principia-qualia/>.
- [9] A. Mabillard. Words Shakespeare Invented, Aug. 2000. URL <http://www.shakespeare-online.com/biography/wordsinvented.html>.
- [10] S. Marshall. Lingua Franca #7: Philosophy. URL <https://mailchi.mp/thestrategicreview/tsr-lingua-franca-7-philosophy>.
- [11] D. Matthews. If you love research, academia may not be for you, Nov. 2018. URL <https://www.timeshighereducation.com/blog/if-you-love-research-academia-may-not-be-you>.
- [12] L. Muehlhauser. Philosophy: A Diseased Discipline, Mar. 2011. URL <https://www.lesswrong.com/posts/FwiPfF8Woe5JrzqEu/philosophy-a-diseased-discipline>.

- [13] A. Pagidas. Philosophy. URL <https://alexandros.is/philosophy/>.
- [14] P. Thiel. We are in a Higher Education Bubble, July 2015. URL <https://www.youtube.com/watch?v=H5NUv0n0QCU>.
- [15] E. Yudkowsky. Evaporative Cooling of Group Beliefs, Dec. 2007. URL <https://www.lesswrong.com/posts/ZQG9cwKbct2LtmL3p/evaporative-cooling-of-group-beliefs>.