Situating LessWrong in contemporary philosophy: An interview with Jon Livengood

117

by **Suspended Reason** 30th Jun 2020

Philosophy | Philosophy of Language | Interviews | Transcripts | Community | Rationality | Curated

Jonathan Livengood is a current associate professor of philosophy at Urbana-Champaign, who hung around the first iteration of LessWrong in the late 2000s and early 2010s as a graduate student at the University of Pittsburgh, where he was writing a dissertation on causal inference under John Norton, Peter Spirtes, and Edouard Machery. He also blogs at the excellent Unshielded Colliders.

One of the central criticisms of mainstream philosophy at LessWrong has always been aimed at its tendency (sometimes called "conceptual analysis") to reify cognitive concepts ir linguistic terms—to perceive them, in other words, as having a simple, one-to-one correspondence with regularities or features of the world (see "Taboo Your Words°," "Concepts Don't Work That Way°," "LessWrong Rationality and Mainstream Philosophy°"). Livengood and I discuss the state of conceptual analysis in philosophy departments, and its recent replacement by "conceptual engineering." We also discuss some of the problems of academic philosophy, continuities between LessWrong and analytic thought, and the status of insights like Bayesianism, verificationism, the pragmatist motto "making beliefs pay rent," and Korzybski's "map and territory."

Some context for this interview can be found in an earlier post, "Conceptual engineering: The revolution in philosophy you've never heard of°," as well as in short pieces on my personal blog about LessWrong vs. contemporary philosophy (1, 2). But I'll add some definitions up front, to give context to our conversation for those who haven't read the backlog:

conceptual analysis: a method of philosophy in which a concept is assumed to have necessary and sufficient criteria which can be described simply and robustly; for instance, there might be a set of criteria which elegantly compress and describe all native-speaker utterances of a concept like "truth." Typically, a philosophical opponent will rebut a proposed set of criteria by offering counterexamples: cases in which a use-case of a concept does not meet the proposed criteria (or in which a non-member of the

conceptual *does* meet them). Michael Bishop's "The Possibility of Conceptual Clarity in Philosophy" is an excellent, if skeptical, introduction.

conceptual engineering°: a recently proposed shift in philosophical method, which abandons the idea of concepts as having "necessary and sufficient" criteria, and instead of analyzing concepts, attempts to rigorize or redefine them so they can be made more useful for a philosophical problem at hand.

This interview runs long, so I've supplied headers and bolded key lines which will hopefully enable selecting browsing.

• • •

On conceptual analysis & the history of philosophy

Livengood: [Before we start, since we're discussing LessWrong versus more traditional philosophy...] It's not clear to me that there's any unique thing we could think of as philosophy, full-stop, or "philosophical discourse today." I think a better picture is there are a bunch of overlapping activities and pursuits; sometimes they have goals that are nearby, and a lot of behavioral practice can live happily in any of those circumstances, but the ends people have in mind are a little different. We can have a lot of shared discourse in philosophical spaces; we all go to the same conferences and there isn't much disconnect, but when you try to get into what exactly people are trying to *do* with these projects, it can come pretty far apart.

Reason: Well, perhaps one angle here—I've heard it argued that conceptual analysis is the foundational, inseparable, aprioristic mode of philosophizing that goes back to antiquity and forms a throughline from philosophy's past to present. And though it's not always stated, the implication is that by turning a leaf from conceptual analysis to conceptual engineering, you've fundamentally changed the nature of the field: what it thinks it's up to in terms of lexicography, how it understands definitions, its place in offering linguistic prescriptions versus descriptions, the factorings of concepts and how people use them, and a larger transition from armchair philosophizing to the kind of experimental, empirical work you're doing with causation. Does that sound like a resonant narrative, or how off am I?

Livengood: I think that's a popular narrative. There's a fair amount of nuance that gets trampled, but it's not a naive or amateurish view, there are philosophers I really like, such

as Stephen Stich, who would give more or less this account of the development of Western philosophy. And you can definitely see elements of it in the Platonic dialogues: Socrates shows up in the marketplace, and someone runs into him, and they say something off the cuff like, "So-and-so was really courageous yesterday," or Euthyphro says, "I'm doing the pious thing by prosecuting my father for murder." And Socrates will go, "*Oh*. So you must know what *courage*, or *piety* is. Tell me about that." The structure usually looks like the other person giving a cluster-type definition, "Piety is when you do these sorts of thing—going to sacrifices, doing what the gods require, visiting the temple on a regular basis." And then Socrates says, "no, I don't want a list."

Reason: He wants the essence.

Livengood: Right, *give me the account*. And the other person realizes what Socrates wants is a definition, so they give an attempt at a definition. Socrates gives a counter-example, so they patch the definition; Socrates gives another counter-example and they patch the definition; and eventually everyone gets tired and leaves. That's the structure of a dialogue, especially the early ones.

There's something really nice about that format, and something that looks very similar to even contemporary work. One of the corners of the literature I know fairly well, the causation literature, a lot of it looks like that. Take David Lewis in the 1970s offering a counterfactual account of causation with a simple core idea: causation is like counterfactual dependence of a certain sort, or some pair of counterfactual dependence claims. And then people point out problems with that account, so he offers patches—in 1986, in 2000, another posthumously. There's a series of counterexamples and revisions to try to capture the counterexamples, and this process repeats and repeats. You wonder if the dialogue's gonna end in the same was as the Platonic dialogues: effectively people get bored with it, and move on, or if there's something like a satisfactory theoretical resolution.

There's an interesting, difficult, subtle kind of question about what the aims of that procedure really are; you'd asked when you wrote me, you used the word "lexicography" in your setup. I don't think for the most part philosophers have been trying to do, or thought of themselves as doing, lexicography. It seems to me that philosophers up until the 20th century, really, were doing one of two things. The boring older thing is doing metaphysics, where the target is supposed to be a thing "out there" in the world, and it's not so much that the project is figuring out how we use language, but about getting at whatever the thing is "out there." Think about this the same way you think about scientific things,

Newton and the apocryphal apple. You say: "That thing we just saw, let's call that gravity; there are objects, and when they're unsupported, they fall." What's the right account of that? We know what we're talking about, we fixed our reference, but now we want to give an account.

It seems to me like historically, philosophers were aiming at the same type of things. You should think of Socrates as saying something like, "We've seen examples of what we might call courage, or piety—there's a *thing*, out there in the world" and here I think he's making a mistake, there's this abstract object "justice" or "piety" or "courage," and that thing I want to give an account of in the same way I give an account of gravity, or matter, or space.

Reason: The mistake being that he reifies a cognitive cluster space of "the good" or "the pious" as matching onto a discernible structure in the world, as opposed to being a garbage heap humans have found useful to call "pious" historically. Do you think philosophy that falls into that style of thought identifies and corrects its mistakes before Wittgenstein, or is Wittgenstein rightfully treated as a big deal in part for noticing it?

Livengood: Wittgenstein is tricky in a few different ways, and the 20th century on this is... contentious. There are two related things that happened where, the history is not so obvious yet, and so there are still live debates about how to think about it. There's this movement of analytic philosophy, you'll see Frege get included, Russell and Moore typically, Wittgenstein and maybe Carnap; sometimes the Ordinary Language group will get picked up like Austin; but there's this core British group that's tough to distinguish from realists.

Reason: They're rebelling from British idealism.

Livengood: And there's this focus on figuring out the meaning of terms; this is a big part of Russell's writing, for example; and there's a lot of concern with the logical structure of speech. Then there's a related phenomenon—sometimes it's smooshed together, sometimes they're separated—this idea of philosophical analysis, and this related idea of the linguistic turn. A number of people think that sometime in the 20th century there's a shift; often they're thinking of Carnap, who is very explicit about the difference between a material kind of discourse, which is how I've described Socrates—giving this account of a thing in the world, like piety—and another mode, Carnap's formal mode, which is, treating this term that shows up in our language, "piety," now with quotation marks. I'm talking about a linguistic object. And of course there's a possible further shift to paying attention to our *concepts*, which are supposed to be attached in some way to a linguistic term.

Reason: I guess one contention I'd advance is, to me, a classical account of concepts as having necessary and sufficient criteria in the analytic mode is in some way indistinguishable from the belief in forms or essences insofar as, even if you separate the human concept from the thing in the world, if you advance that the human concept has a low-entropy structure which can be described elegantly and robustly, you're essentially also saying there's a real structure in the world which goes with it. If you can define X, Y, & Z criteria, you have a *pattern*, and those analyses assume, if you can describe a concept in a non-messy way, as having regularity, then you're granting a certain Platonic reality to the concept; the pattern of regularity is a feature of the world. I don't know, what do you think of that?

Livengood: There's a lot right about what you said, and the kinds of challenges you see in the middle of the 20th century are serious problems for this whole collection of approaches, but I think it's important to see that this kind of move, especially from Carnap, which was prefigured a bit by what Russell was doing, was an important advance because it didn't necessary reify the target of the inquiry. In some cases you might want to say, "Gravity, that's something we can responsibly talk about as existing in the world," but for other things, we might just want to talk about what our language is doing. It might just be transactional—what kind of inferences we're going to make, what linguistic acts we're gonna trade back and forth; it might not be tracking anything out in the world. So there's been a pretty serious advance from the picture you're getting from Socrates up through the 20th century, to when people start focusing on the language, and thinking of linguistic acts or the structure of the language as themselves the targets of the investigation.

Reason: It's hard to understand the history backwards; much of what past philosophers got right now seems obvious, while everything non-obvious is wrong.

Livengood: I think that's right; **one of the things that's fun about doing history of philosophy is seeing how very smart people can be deeply confused about things**. They have an *idea* but it's vague and mashed-up, and today you'd say, "You're running together six different things, you have to pull apart and distinguish them." It's a thing that happens a lot, reading the history.

Reason: If I want to learn about the history of philosophy, or what Kant thought, or about philosophy *through* Kant—in which of these situations should I read the original, and when should I read a secondary source?

Livengood: Secondary sources have huge virtues, and you've identified some of them: they're often clearer than primary sources, they often supply intellectual context and help

situate the primary source, while drawing out what the field thinks is important. But there are also vices: the secondary literature may not be right about what the most important things in the primary source are; often these sources are idiosyncratic in their readings.

Reason: What's your gut on how good these secondary sources are? Let's say major university press, respected in the field. Have we pretty much mined everything in the original, or are there gems still hiding out?

Livengood: **The danger is more on the side of over-interpreting, or being overly charitable to the target**. I just wrapped up a grad seminar on the problem of induction, and we were looking at the historical development of the problem of induction from Hume to 1970. As I pointed out, when you look at Hume, Hume's great, he's fun to read, but he's also deeply confused, and you don't want to do the following, which is a mistake: If you start with the assumption that Hume was just *right*, and assume that, if you're seeing an error it must be an error in your interpretation—if that's your historiographical approach, you're not going to understand Hume, you're going to understand this distorted SuperHume, who knows all these things Hume didn't know, and can respond to subtle distinctions and complaints that someone living now is able to formulate. That's not Hume! Hume didn't have an atomic theory, he didn't know anything about DNA or evolution; there are tons of things that were not on his radar. He's not making distinctions we'd want him to make, that a competent philosopher today would make. There's a real danger writing secondary literature, or generating new interpretations. **If you want to publish a book on Hume, you need to say something new, a new angle—what's new and also responsible to what Hume wrote?** It ends up doing new philosophy under the guise of history. There I'm suspicious that there's anything new to say that's also responsible to the writer.

In the 70s, the target for me is Quine; he wrote a paper called "Epistemology Naturalized," and there's a straightforward reading of this paper where he's resuscitating Hume, and giving a contemporary update. He has this throw-away line; the slogan part is, "The Humean predicament is the human predicament," but he also says, *there hasn't been any progress in epistemology on the doctrinal side, the side that's dealing with normative questions, questions of justification, and the problem of inductive reasoning, since Hume*. So the seminar [I ran] was asking: Is Quine right? I was upfront with the students, that there's been a lot of work on inferential problems between 1970 and today; almost all the interesting work on causal inference is after 1970. You have the emergence of information criteria, lots of statistical techniques like the bootstrap and jackknife, Bayesian and computational resources, machine learning and big data—those all change the landscape.

Livengood's experience with LessWrong

Reason: I want to ask how you think of the historic state of philosophy, or what it would be like to project a historical view on the present, but I want to ask about LessWrong, so let's jump back and forth. How'd you get exposed to the community? What was your experience?

Livengood: I started reading in the 2000s, I don't remember exactly which pieces. Much of it was just self-reinforcing; for the most part, stuff that happened on LessWrong [then] seemed indistinguishable to me from high-level amateur, low-level professional discourse in philosophy? Smart graduate students, people who had really decent ideas but lacked the professional language to express it. That's the way the LessWrong community struck me at the time; I was a graduate student just starting, and it felt like, "Yeah! I'm having a conversation with other people doing the same kind of thing I'm doing." **There's sometimes an impression that the people on LessWrong were doing something wildly out of step from what philosophers would ordinarily think of themselves as doing, and that was not my impression.**

Reason: Both naysayers and advocates for LessWrong or Yudkowsky do often emphasize the gap like you say, and I think unless you're very knowledgeable about the field, you hear a lot of bad arguments coming out of philosophy, both historically and still today. (Sturgeon's Law.) And most philosophers worth their chops in these fields are aware of these historical arguments being flawed; they're maybe more generous, and probably see these (today obvious) ideas as highly non-obvious in their times.

Livengood: Again, the thing I said earlier, that there isn't "such a thing, fullstop" as philosophy—LessWrong [at that time] seemed fruitfully engaged in similar kinds of questions, concerns, and problems to at least some parts of contemporary academic philosophy, and parts of contemporary philosophy I like and think are non-trivial. It's not a ghettoized, small corner of philosophy; there are robust projects that are shared by a number of departments across the world that do things this way.

I would agree LessWrong does things differently, there's a house style, but it's not like the collection of theses they defend or are pursuing or developing are so far out of the mainstream that academics wouldn't recognize it as philosophy, or as being reasonable approaches to philosophy.

Romantic vs. professionalized philosophy

Reason: Well, that's why I reached out in the first place; you'd left a comment° on Luke Muehlhauser's "Train Philosophers With Pearl and Kahneman, not Plato and Kant°" gesturing to this effect—that at least in your graduate program, at Pittsburgh, cognitive science was very paid-attention-to.

Livengood: The Pittsburgh scene is a little peculiar; just background-wise, at the University of Pittsburgh there are two departments which at the time were on the same floor. There's an enormous, 42-story cathedral of learning at Pittsburgh, lovely neo-Gothic, built in the 30s, and these two departments were right across the hall: there was the philosophy department, and there was the History and Philosophy of Science (HPS) department. My PhD is from the latter.

Those departments are very different in the way they think about what philosophy is doing, the way they train their graduate students, the way their courses are conducted, their faculty. Maybe the best way to describe that difference is there are two divergent attitudes of how philosophy should go, what I'd describe as the professionalized view and the romantic view. The HPS side tended to be more professionalized; you find an interesting problem, chip away at it, advance the field a bit, and at the end of a long career, you and the people you're working in conversation with will have learned something, you'll have advanced human knowledge. This is the way things have to go: most of us are not geniuses, we're just ordinary people chipping away at a problem.

And then there's the romantic view that says look, the people we read and engage with—Aristotle, Descartes, Kant, Wittgenstein—are these super-geniuses who thought thoughts nobody else had ever thought before, who shook the foundations of human knowledge and turned things upside down. This is the aim: to become one of those people.

And the difference in graduate training in the two programs is, HPS you come in, write some papers, get out in 6-8 years, get a job, everybody does that. The Pitt Philosophy program you come, think some things, try to think the deep thoughts; the very best people go on to an awesome career, the rest of you, well, we're happy to burn through a hundred grad students to find a diamond.

My sympathies are, as you might expect, entirely with the professionalized view.

Reason: It does seem if you're a Wittgenstein-level genius, you don't need your romanticism stoked, you might not even a graduate program. Certainly they didn't.

Livengood: That's probably right, but to give the devil his due, there are things to like, there are reasons people are attracted to that romantic view. They're just not reasons I endorse at the end of the day.

Analytic communities on LessWrong's wavelength

Reason: Have you read Clark Glymour's manifesto°?

Livengood: Yes.

Reason: What did you think?

Livengood: So that's the other element in the mix. There are these two Pitt departments, both quite good, the Philosophy program at the time was top five in the world, and HPS program has been for a long time *the place* to do philosophy of science. And then across the street is Carnegie Mellon, which, their philosophy department is basically Glymour's construction. Whoever the president or provost was recruited Clark out of Pitt to establish a philosophy department, and Glymour's like, great, I can build a philosophy department from scratch, the way I'd want to run a philosophy department. It's a peculiar place. **The way I've heard it described is that CMU's philosophy department is what you get when you treat philosophy as a kind of engineering.** I think that's not inaccurate. I happen to think that's beautiful, a really good look for philosophy.

Reason: What would you call the CMU, HPS, maybe LSE, you can throw LessWrong in there it sounds like—

Livengood: I would include also Irvine, University of Minnesota, Indiana University sometimes has had this vibe. It's not quite positivist, but it's in that neighborhood—science-friendly, professionalized, trying to make progress, caring about mathematics and empiricism.

Reason: It's the kind of people who would've been positivists in the 50s.

Livengood: If Carnap were alive today he'd be in this camp. Whether he'd have the views he had back then, well, he probably wouldn't; we learn things, we hope that these things change minds.

Reason: I've heard this vibe is also popular in Europe.

Livengood: Yeah, the LMU at Munich has the same kind of character. European programs are trickier because much of it is tied to local funding regimes, but there do seem to be more of these mathematically, empirically informed projects.

Reason: A popular metaphor at LessWrong is Korzybski's "map and the territory," though it may have gotten there via Hayakawa. Is it a good metaphor, or do its reductions actually set you back, as some detractors claim?

Livengood: I think I'm mostly a fan of the Korzybski metaphor. It's serviceable. I think it has some limitations where the map *is* the territory, which can happen when the map-making makes the thing. Here I'm thinking of pretty mundane cases, like how something being *money* depends on how we treat it, and also more controversial cases, like the construction of gender and race or the status of mathematical objects. Or do you think that misses the point of the metaphor?

Reason: Bayes, underrated, overrated?

Livengood: Hm... a bit of both. Bayesian approaches in philosophy of science and epistemology today are pretty standard. Bayesian analysis of scientific reasoning is a project that's probably overrated, at least in philosophy. Bayes in undergraduate education generally is probably underrated; I teach a 100-level intro to logic course, and I tell the students, if you take a Stats 100 class, you'll see frequentist approaches to probability, and frequentist statistical inference techniques, so I'm going to give you something different, give you a Bayesian take on it. So far I haven't yet have a student saying, well, this is obviously the way people think about probability, this is boring and I've seen it in my other classes.

Reason: We're obviously familiar with the idea of scientific progress. Ethics get described surprisingly similarly, where there's a kind of drift; whether that drift happens "on its own," in an inevitable ratchet, or whether people have to work to make it happen, is unclear; but this is the way changing norms around race, sexuality, animal rights get talked about typically. Do you feel like the shift that departments like HPS or CMU are leading, the transition from conceptual analysis, will win out or become dominant? How do you see the field a hundred years out?

Livengood: Predictions that far out are tricky. It's not obvious to me we'll have anything that look like contemporary universities in a hundred years. You asked over email about technological developments and philosophical progress, and there are lots of positive

impacts there. Increases in massive online instruction, I'm not sure how that will shake out.

Philosophy's role in public discourse

Reason: Last year you wrote, "I don't think philosophers are especially well-equipped in virtue of their training to help out in the current crisis. We're more like high-trained sports fencers when a general melee is breaking out. We've trained to participate in a game that has specific restricted rules, that are implicit and often hard to fathom; if we go out into the world and try to fix it playing by our usual rules, the result will be predictably bad." This seems right to me, but the question becomes, who is filling this role? We don't have literal swordfights, so it's not a big deal if human capital is channeled into play-fencing. We do have these figurative swordfights though, so the question becomes, who is filling this role in public discourse?

Livengood: I thought your list was pretty good. [*I'd emailed along Tyler Cowen's comments that amateurs in philosophy are running the public-facing discipline: Silicon Valley stoicism, Nicholas Nassim Taleb, LessWrong-style rationalism and post-rationalism, ex-New Atheists like Sam Harris, psychologists like Jordan Peterson.*] It gets filled in a variety of a way, some by professional or near-professional philosophers by way of podcasts, but much of it in larger circuits are indeed filled by people like Sam Harris, Jordan Peterson, and then even less interesting people like Ben Shapiro.

Reason: Zizek seems like one of the few entries from a more traditional philosophy tradition.

Livengood: Yeah, there are a few outliers. Peter Singer has had a fair amount of popular public impact. There are other with marginal public influence, but who are clearly important, such as Martha Nussbaum or Dan Dennett. They matter, even if they're not nearly as visible as people like Zizek, or Chomsky, or Singer. I don't know how many public-facing philosophers we need in a society of this size; it does seem like, given that I'm not especially impressed by people like Harris and Peterson and Shapiro, we could use more public-facing philosophy—but **there's also a question of why it is the market has taken up those individuals, whether there are just market-type demands that are satisfied by the ideas they're producing that wouldn't take up public bandwidth the way more mainline philosophical production would**.

Reason: Looking to one historical precedent, what do you think of say the post-war French gang, Sartre through Foucault? That's a case of borderline public hysteria around a set of more-or-less traditional academic philosophers. Is that fair? What can we take away, what do we learn?

Livengood: I'm not sure we learn anything. I'm not a radical contingency historian, I don't think there's nothing to learn from history, but there are often events where there isn't much to take away, you have a couple interesting public intellectual figures who happen to be in philosophy, who happen to have a public who is interested in their ideas; if they'd been in a different field, would things have been different? I don't know. The counterfactuals make me think it's too hard to judge. At minimum, we'd need a whole lot more detailed information about their writing, what was going on in society, and I'm unqualified for that.

Selection and referee problems in philosophy

Reason: I've really appreciated how much personality philosophy has. You have Chalmers and his Zombie Blues band, the Kripkensteins, it's a fun wonky field, old men with big personalities and big beards, I'm a big fan. But now that I've said something nice about philosophy I have to say something mean. Sturgeon's Law says 90% of any field is bad, 10% is good; you have plenty of dressed-up, garbage literary fiction and plenty of brilliant pulpy sci-fi books. Do you think there's a mechanism that makes it more difficult for the field to sort out and identify the good among the bad? Maybe it takes a certain level of criticality to identify the good thought to begin with, and the implicit consensuses built off support and textual elaboration aren't guiding us to the correct answers.

Livengood: Part of what you're saying sounds right, but I'm a little nervous about other bits. I'd put it in terms of "rules for settling opinions": in the sciences, there are clear standards for settling disputes, where you work out an experiment and run it. I'm not naive about how the sciences work in reality, but in principle at least, if you have a disagreement, you can come to an agreement about what you will do or believe in light of the experiment you're going to run. This is an idealistic Feynman picture, that at the end of the day, if you run the experiment, and the experiment doesn't agree with your idea, even if your idea is super pretty, it's wrong. In real scientific practice it's a lot messier, but in philosophy it's much harder to agree on a constraint or rule for settling disputes. We have practices we engage in, and we do tend to move closer together in the process of extended discourse and argument, but it's hard to say why that happens; I find it very

unsettling that I don't a good sense of what might resolve a disagreement. It's a problem I'm always puzzling about.

There's something I want to fuss about though. It seems to me that philosophy has a bad cultural fixation on the genius, but that a lot of progress is possible in philosophy without these super-genius-level contributors. This is part of my bias toward the professionalized way of looking at the field. I think the best work in philosophy is identifying a narrow topic you can actually make progress on, and chipping away at it through formal precision, distinctions, experiments, and collectively we make progress on these problems. It's not always obvious that there's progress, or what progress looks like, when you're too close to it, or it's really new, but if you give yourself an extended period—how people have thought about induction from Hume to today—you'll see *lots* of progress made.

Making beliefs pay rent

Reason: I can't let you go before asking about Peirce, who you've written quite a bit about. One of the views of his that surfaces on LessWrong is a demand that beliefs pay rent. Now, I know people make a lot of the differences between pragmatism and positivism, and certainly Russell hated the pragmatists, but there seems to be a kernel or core, maybe you could call it weak verificationism, where if one person believes one thing, and another believes another thing, then there should be some observable difference that matters, something that ought to tell us who is right our wrong. That if there's nothing in the world that can distinguish between our arguments, maybe we're not in disagreement at all. Verificationism proper comes under a lot of flack these days; maybe you can suggest a better handle for the rough, generic version I'm describing; but I'm curious, is verificationism a good idea that's needed a lot of qualification over the 20th C, or is it a bad idea that got us off on the wrong foot?

Livengood: I think it's a great idea that's mostly right. It's similar to what we were talking about with primary and secondary sources: the bulk of its value lies in pretty simple statements, even though those statements aren't quite right. They have counterexamples, or haven't had enough detail built into them, but you get the gist. **It's still an open question as to whether an adequate account of the verification criterion can be made to work, but I'm not sure it really matters with respect to the practical service the idea performs.** Something like Peirce's pragmatic maxim, or various Positivist views, or the verificationism Quine goes in for—all of those are quite salutary attitudes to have.

Broadly good, broadly healthy, and they inspire broadly good practices in our intellectual lives.

Now, when you start trying to narrow it down to a dogmatic thesis, then I'm not so sure a verificationist account of meaning is going to quite work. There are some obvious failures; A.J. Ayers' account doesn't work, it's pretty easy to kill it, and Church gives devastating counterexamples.

Reason: If we cast Ayers as a conceptual engineer, isn't he just telling us what a meaningful sentence is?

Livengood: Yes! This is more or less the Carnapian route. Carnap's accounts have not been knocked over in the way Ayers has been.

Reason: Well, I'll just ask a couple minutes more of your time: One paper I've gotten a lot out of is Michael Bishop's 1992, "The Possibility of Conceptual Clarity in Philosophy." He talks about a "counterexample" style of philosophizing that's broader than conceptual analysis, where the philosopher sits in the figurative armchair, proposes a definition, and another armchair-occupant posits a counterexample which pokes a hole in the original proposal. Much like a Socratic dialogue. Given this has been the standard method for both proposing and rejecting proposals, it seems that, if we grant prototype theory and reject classical accounts of concept—if we believe concepts are fuzzy and polysemous; that there will always be edge-cases to a conceptual carving, and there's no way to losslessly compress into a few simple criteria the high entropy use-in-the-world by millions of decentralized speakers over time—if we grant this about concepts, should we let the classically analytic rulings from the 20th C about what is "meaningful" or "true" or "knowledge" stand? Ought we revisit those debates to see if they might be useful factorings, even if they aren't necessary and sufficient?

Livengood: Yes. The best example I can give is work by Joseph Halpern, a computer scientist at Cornell. He's got a couple really interesting books, one on knowledge one on causation, and big parts of what he's doing are informed by the long history of conceptual analysis. He'll go through the puzzles, show a formalization, but then does a further thing, which philosophers need to take very seriously and should do more often. He says, look, I have this core idea, but to deploy it I need to know the problem domain. The shape of the problem domain may put additional constraints on the mathematical, precise version of the concept. I might need to tweak the core idea in a way that makes it look unusual, relative to ordinary language, so that it can excel in the problem domain. And **you can see how he's making use of this long history of case-based, conceptual analysis-friendly**

approach, and also the pragmatist twist: that you need to be thinking relative to a problem, you need to have a constraint which you can optimize for, and this tells you what it means to have a right or wrong answer to a question. It's not so much free-form fitting of intuitions, built from ordinary language, but the solving of a specific problem.

Philosophy 3 | Philosophy of Language 1 | Interviews 1 | Transcripts 1 | Community 4 | Rationality 2 | Curated

Mentioned in

74 2020 Review Article

22 comments, sorted by top scoring

[–] **Charlie Steiner** 3y 24

I've actually been recreationally looking at some undergraduate philosophy courses recently. And it still shocks me just how backwards-looking it all is. Basically nothing is taught as itself - it's only taught as a history of itself.

There are two main skills that I think are necessary to practice philosophy (at least the sort that I have practical use for): the ability to suspect that your model of things is wrong even as you try your best, and the ability to sometimes notice mistakes after you make them and go back to try again.

Presumably this is what grad school is for, in one's philosophy education, because I haven't seen deliberate practice of either in the lectures and books I've skimmed. The presence of this sort of thing is one of the factors that makes LW stand out to me. If one were situating it not within philosophy but within philosophy education, it would be a pretty nuts outlier.

[–] **Stabilizer** 3y 5

If you'd like to learn non-backwards-looking philosophy, which is indeed how most philosophy in mainstream American departments is done, then I highly recommend skipping undergraduate courses, which for some weird reason, kinda "talk down" to the students. Instead, I suggest three things:

(1) Just read the Stanford Encyclopedia of Philosophy. Pick a topic you like, such as causation or time or animal ethics, and just read the article or related articles.

(2) Read or skim academic papers or books. Most of them are surprisingly readable, especially the introductory parts. Notwithstanding criticisms of academic writing, I do think that analytic philosophy places unusual emphasis on writing clearly and plainly. (We can thank Russell and Moore for that in large part. Though, Plato wrote beautifully as well.) You can find good ideas for what to read from the Stanford Encyclopedia or track down philosophers whose work you find interesting.

(3) Listen to podcasts. Philosophy Bites's archive is a treasure trove: it has so many important philosophers on and they all have interesting and clear explanations of some central idea. Also check out Matt Teichman's elucidations. And there are a few more I'm forgetting. And then if you find something interesting, track the

philosopher down, and read their books or papers. (Unfortunately, blogging by philosophers isn't as active as one might wish; I think this tracks the general reduction in blogging on the Internet.)

You'll learn a lot more this way than through undergraduate classes, which are usually slow and dull. I'm in philosophy grad school, but never took any philosophy undergraduate classes, but I picked up a significant background in philosophy using the 3 techniques above. I'm really happy for that. I love research-level philosophy, but undergraduate classes are too slow for me to sit through.

[−] **TAG** 3y 2

> The ability to suspect that your model of things is wrong even as you try your best, and the ability to sometimes notice mistakes after you make them and go back to try again.

Wouldn't knowledge of past mistakes be helpful in that? Wouldn't recreating philosophy from a blank slate lead to reproducing a lot of errors?

[−] **Charlie Steiner** 3y 5

Yeah, if there was a philosophy course somewhere that actually looked at past mistakes in philosophy, that would be really interesting.

But on the other hand, doing that gives up some sort of cosmopolitanism that I think university philosophy courses really try to hold on to. Even if you take a philosopher with a strong opinion on the subject, when it comes to teaching a course they'll probably teach a section on Kant that faithfully explores the history of Kant's arguments about analytic vs. synthetic, and then a month later they'll get to Quine and faithfully explore the history of Quine's arguments about analytic vs. synthetic.

Now, one might say "Those people really disagreed with each other, so clearly at least one of them has made a mistake. Why are you just repeating what their arguments were rather than helping the students actually learn how to avoid the mistakes of the past?"

I think partly it's covering your ass against bias, real or merely accused. If experts disagree about something, then as a teacher you should probably teach both sides to some extent, even if you have an opinion.

But I think the real reason, which happens to be a worse reason, is that the knowledge of historical arguments is what people see as the core of a philosophy degree, and ability to avoid the mistakes of the past is not as central. If some department full of Quineans just totally skipped over the notion of analytic vs. synthetic in their courses because they didn't think it's a useful distinction, there would be this sense of "well, they haven't *really* gotten a philosophy education" that stems from associating a philosophy education with trivia rather than skills.

[−] **TAG** 3y 0

In a sense it's all about mistakes ,because the history of philosophy isn't a bunch of random stuff, it's one phlosopher reacting to another.

But you seem to want mistakes in a sense where they are not just criticisms from some perspective or set of assumptions, but absolute. That you are not going to get , because epistemology has not been solved. So what you have instead is everyone criticising everyone else in a Mexican stand off.

If it were possible to divide philosophy into right stuff and wrong stuff ,you would need an explanation, such as cosmopolitanism, for continuing to teach the wrong stuff. But that would be downstream of solving episyemology.

[–] **Charlie Steiner** 3y 2

I broadly agree.

I think we'd agree that *some* philosophical progress has happened over the last couple thousand years, though (though I'd probably claim there's been a lot more progress in epistemology since 1950 than you'd agree with). Our hypothetical "Mistakes of the past" philosophy course couldn't just be a regular survey course but with the professor taking sides on every issue, but it could be cherry-picked to take advantage of places where the issue appears clear-cut in hindsight.

Since you can find someone to disagree with anything, of course for each mistake you could find someone who disagrees, so the amount of editorial control isn't zero, but in general I think that this kind of material would actually be appropriate for a liberal-arts setting. u/Jonathan_Livengood you should get on developing this course :P

[–] **Jonathan Livengood** 3y 3

It's a very interesting suggestion. I haven't really taught history of philosophy --- a big exception being a graduate course on the history of work on the problem of induction from Hume to Quine, which I taught in the spring. Basically all of the courses I've taught are current topics, arguments, and controversies that are live today. Course titles like "Logic and Reasoning," "Biomedical Ethics," "Contemporary Philosophy of Science," "Metaphysics," and "Philosophy of Psychology."

Maybe the way to teach something like this would be under the heading "Progress in Philosophy," where you could sort of split time between [1] the contemporary debate about what counts as progress and whether there is or could be progress in philosophy and [2] some historical examples. (This was also a major theme of the grad course I taught in the spring, so it's still very much on my mind.)

[–] **Charlie Steiner** 3y 2

Out of curiosity I looked up what you were teaching in the spring - the problem of induction, right? (I'll be surprised and impressed if you managed to foist a reading from *Li and Vitanyi* on your students :P) I'm definitely curious about what you think of the progress in probability, and what morals one could draw from it.

I'd actually checked because I thought it would be the philosophy of psychology. That seems like one of those areas where there were, in hindsight, obvious past mistakes, and it's not clear how much of the progress has been emprirical versus things that could have been figured out using the empirical knowledge of the time.

[–] **Jonathan Livengood** 3y 1

No Kolmogorov complexity -- the course was really a history from Hume to about 1970. The next time I teach a seminar, I'm hoping to cover 1970 to the present. Still, this time around, a lot of the readings were technical: Ramsey, Jeffreys, Fisher, Neyman, De Finetti, Savage, Carnap, and others. You can see the full reading list here.

I agree that a nice course on progress could be done with a philosophy of psychology focus. I expect that progress-skeptics would object that the progress is in psychology itself, not in the philosophy of psychology. (I wouldn't share that skepticism for a couple of reasons.) Maybe if the course were framed more in terms of philosophy of mind and computation? Have you read Glymour's "introduction" to philosophy, *Thinking Things Through?* It has that feel to me, though it's pitched more like, "Here are things that philosophy has contributed to human knowledge," and it ranges over more than mind and computation.

[−] **TAG** 3y 1

> it could be cherry-picked to take advantage of places where the issue appears clear-cut in hindsight.

If you did that 1500 years ago, then theism would appear clear cut in hindsight.

If you did that 150 years ago, then reductionism would appear obviously false.

> this kind of material would actually be appropriate for a liberal-arts setting

As opposed to what? Would you be doing anyone any favours by rounding off "seems true to us, here now" as the last word on the subject?

[−] **Charlie Steiner** 3y 2

Yes. Favors would be done.

I'm actually not sure what argument you're implying by your past examples. 1500 years ago the denial of Euclid's parallel postulate wouldn't have been taught - does this have implications for modern mathematics education?

[−] **TAG** 3y 3

It has implications for physics. If you re run the history of thought with even more emphasis on what's currently believed to be true,and even more rejection of alternatives, then you just slow down the acceptance of revolutionary ideas like non Euclidean geometry.

> Favors would be done

Would they? Can you explain how and why?

[−] **Charlie Steiner** 3y 3

But past mathematicians already just taught what they thought was true then. I'm not asking why they didn't do that even harder, I'm asking what relevance you think it has for current math education. (And by extension, what relevance you think the education system of the Scholastics has for modern philosophy education.)

As it is said, keep an open mind, but not so open your brain falls out. Teaching a specific thing impedes progress when that thing is wrong or useless, but it aids progress when that thing is a foundation for later good things. This framework largely excuses past mathematicians, and also lets us convert

between the "cautiousness" of philosophy education and a parameter of optimism about the possibility of progress.

[–] **TAG** 3y

> But past mathematicians already just taught what they thought was true then.

But we don't know that we are living in the optimal timeline. Maybe relativity would have arrived sooner with fewer people in the past insisting that space is necessarily Euclidean.

> I'm asking what relevance you think it has for current math education.

The topic is philosophy education. Science can test its theories empirically. Philosophy can't. Mathematics can take its axioms for granted. Philosophy can't.

> As it is said, keep an open mind, but not so open your brain falls out. Teaching a specific thing impedes progress when that thing is wrong or useless, but it aids progress when that thing is a foundation for later good things.

The difficulty is that we don't have certain knowledge of what is in fact right or wrong: we have to use something like popularity or consensus as a substitute for "right".

It may well be the case that one can go too far in teaching unpopular ideas, but it doesn't follow that the optimal approach is to teach only "right" ideas, because that means teaching only the current consensus, and the consensus sometimes needs to be overthrown.

The optimal point is usually not an extreme, or otherwise easy to find.

[+] [comment deleted] 3y

Deleted by TAG, 07/03/2020
Reason: Comment deleted by its author.

[–] **Jayson_Virissimo** 3y

Suspended Reason: you may find this° philosophy poll of LWers from 8 years ago interesting. The poll results no longer render (as of the 2.0 reboot of LW), but the raw data can be found in this git repo.

[–] **Suspended Reason** 3y

Thank you! I'd seen the poll but not the repo.

[–] **Jayson_Virissimo** 3y 4

Good call, I'll link to it from the poll.

[–] **habryka** 3y 7

Promoted to curated: I really enjoyed reading it, mostly because it helped me put a bunch of knowledge from LessWrong into a broader context. I also disagree with some good chunk of the post and some of the analogies, but don't think that really makes it worse, in many ways I am glad to have a perspective presented that feels somewhat alien to me, and therefore something that might have insights that I am particularly likely to miss.

Overall, I was quite glad to have read this, and am really glad about this interview happening.

[–] **Suspended Reason** 3y 1

Thank you! I'd be very curious to hear what didn't resonate, since I'm working the ongoing MetaSequences project, but of course you're very busy, so only if you think it'd be valuable for both of us!

[–] **Chris_Leong** 3y 4

"And the difference in graduate training in the two programs is, HPS you come in, write some papers, get out in 6-8 years, get a job, everybody does that. The Pitt Philosophy program you come, think some things, try to think the deep thoughts; the very best people go on to an awesome career, the rest of you, well, we're happy to burn through a hundred grad students to find a diamond." - I found this passage surprising. I'd expect that the ease of finding a job in an area such as philosophy or HPS would be based on the availability of funding, not differences in approach.

[–] **ESRogs** 3y 2

Are you imagining them competing for two different pools of funding?

Moderation Log