# Why Isn't There More Progress in Philosophy?[1]

DAVID J. CHALMERS

**Abstract**
Is there progress in philosophy? A glass-half-full view is that there is some progress in philosophy. A glass-half-empty view is that there is not as much as we would like. I articulate a version of the glass-half-empty view, argue for it, and then address the crucial question of what explains it.

Is there progress in philosophy? I have two reactions to this question. First, the answer is obviously yes. Second, it is the wrong question. The right question is not 'Is there progress?' but 'Why isn't there more?'.

We can distinguish three questions about philosophical progress. The Existence Question: is there progress in philosophy? The Comparison Question: is there as much progress in philosophy as in science? The Explanation Question (which tends to presuppose a negative answer to at least one of these two questions): why isn't there more progress in philosophy?

What we might call a glass-half-full view of philosophical progress is that there is some progress in philosophy. The glass-half-empty view is that there is not as much as we would like. In effect, the glass-half-full view consists in a positive answer to the Existence Question, while the glass-half-empty view (or at least one salient version of it) consists in a negative answer to the Comparison Question. These views fall between the extremes of a glass-empty view which answers no to the Existence Question, saying there is no progress in philosophy, and a glass-full thesis which answers yes

[1] I first gave a brief version of this paper at the Harvard-Australia conference on Progress in Philosophy at Harvard University in 2011. I thank that audience and those on subsequent occasions in Arizona, Cambridge, Fordham, Liverpool, Oslo, Rio, Santiago, and at the Royal Institute for Philosophy lecture in London. I am also grateful to participants in a number of useful Internet discussions. For comments on the written version, thanks to Russell Blackford, Melissa Ebbers, Alan Hájek, Robin Hanson, John Keller, Mark Lance, Seth Lazar, Christian List, Luke Muehlhauser, Rick Repetti, and Joshua Weisberg.

doi:10.1017/S0031819114000436 
First published online 3 November 2014

to the Comparison Question, saying there is as much progress in philosophy as in science (or as much as we would like).

Of course the glass-half-full thesis and the glass-half-empty thesis are consistent with one another. I think for almost anyone deeply involved with the practice of philosophy, both theses will ring true. In discussions of progress in philosophy, my experience is that most people focus on the Existence Question: pessimists about philosophical progress[2] argue for the glass-empty thesis, and optimists[3] respond by defending the glass-half-full thesis. I will focus instead on the Comparison and Explanation Questions. I will articulate a version of the glass-half-empty thesis, argue for it, and then address the crucial question of what explains it.

I should say that this paper is as much an exercise in the sociology of philosophy as in philosophy. For the most part I have abstracted away from my own philosophical and metaphilosophical views in order to take an 'outside view' of philosophical progress from a sociological perspective. For much of the paper I am largely saying the obvious, but sometimes the obvious is worth saying so that less obvious things can be said from there. Only toward the end will I bring in my own views, which lean a little more toward the optimistic, and see how the question of philosophical progress stands in light of them.

## 1. The central thesis

The form of a glass-half-empty thesis is: there is less progress in philosophy than some benchmark. To articulate such a thesis more precisely, one needs to articulate a measure of progress and a benchmark. The measure of progress I will use is collective convergence to the truth. The benchmark I will use is comparison to the hard sciences.

Here I take inspiration from Peter van Inwagen, who writes:

> Disagreement in philosophy is pervasive and irresoluble. There is almost no thesis in philosophy about which philosophers agree. If there is any philosophical thesis that all or most philosophers affirm, it is a negative thesis: that formalism is not the right

[2] E.g. Dietrich, E. 'There is no progress in philosophy', *Essays in Philosophy* **12** (2010): 329–44; Nielsen, K. 'Can there be progress in philosophy?', *Metaphilosophy* **18** (1987): 1–30; McGinn, C. *Problems in Philosophy* (Oxford University Press, 1993).

[3] E.g. Stoljar, D. *Philosophical Progress: In Defense of a Reasonable Optimism* (forthcoming).

> philosophy of mathematics, for example, or that knowledge is not (simply) justified, true belief.
>
> That is not how things are in the physical sciences. I concede that the "cutting edge" of elementary-particle physics looks a lot like philosophy in point of pervasive and fundamental disagreement among its respected practitioners. But there is in physics a large body of settled, usable, uncontroversial theory and of measurements known to be accurate within limits that have been specified. The cutting edge of philosophy, however, is pretty much the whole of it.[4]

Van Inwagen's thesis is not explicitly about progress, and the general tenor of his discussion suggests something closer to a glass-empty thesis than a glass-half-empty thesis. I think that at least once the issue is made a little more precise, however, a glass-half-empty thesis is more defensible.

Here is my central thesis: There has not been large collective convergence to the truth on the big questions of philosophy.

Here the big questions of philosophy are questions like: What is the relationship between mind and body? How do we know about the external world? What are the fundamental principles of morality? Is there a god? Do we have free will? I will not try to provide a more precise list than this, but any philosopher can come up with a list of 10 or so big questions fairly easily, and I suspect that there would be a lot of overlap between these lists. We could even use these lists to operationally define the big questions: the big questions of a field at time $t$ are those that members of that field would count as the big questions of the field at time $t$. For purposes of comparison, we may want to impose some regimentation on the form of the big questions, for example formulating them all as choices between a small number of mutually exhaustive options.

We can define collective convergence on an answer over a period as the increase in degree of agreement on that answer from the start of the period to the end of the period. Degree of agreement can be defined using one of various mathematical measures of agreement across a group of people on a set of issues.[5] Collective convergence

[4] Van Inwagen, P. 'Freedom to break the laws' in *Midwest Studies in Philosophy* **28** (2004): 332.

[5] I leave open the question of just what measure of agreement is best for present purposes. One useful measure is Krippendorff's alpha (Krippendorff, K. *Content Analysis: An Introduction to its Methodology*, 3rd edition (Sage Publishing, 2013), 221–50), equal to $1-(D_o/D_e)$, where $D_o$ is the observed incidence of disagreement between respondents



https://doi.org/10.1017/S0031819114000436 Published online by Cambridge University Press

(simpliciter) over a period is defined as the collective convergence on the dominant answer at the end of that period over the period.

Collective convergence on the big questions over a period is stipulated to be convergence on the questions that were big at the start of the period. What matters here is convergence on the big questions of the past rather than on big questions of the present. Convergence on the latter can be expected to be low, as disagreement on the answer to a question is highly relevant to making it a big question. Convergence on the former provides a more discriminating measure which can reasonably be expected to be high in some cases and low in others. For our purposes we could choose an arbitrary period (say, from 200 years ago to today), or perhaps better, choose many different periods and take the average convergence over those periods.

We can say that *large* collective convergence over a period requires as much convergence as there has been over big questions in the hard sciences in the same period. Here I will take the hard sciences to include at least mathematics and the natural sciences: paradigmatically physics, chemistry, and biology. I set aside the cognitive and social sciences, which have arguably seen less convergence on the answers to their big questions. For current purposes I do not need to take a stand on how philosophy fares relative to these.

Large collective convergence to the truth in a period requires large collective convergence to *correct* answers to the big questions over that period. That is, we must have as much increased agreement on correct answers to the big questions as in the hard sciences. Because of the reference to truth or correctness, large collective convergence to the truth requires a degree of realism about the domains in question.

---

(summed over all pairs of respondents and all questions) and $D_e$ is the expected incidence through chance alone. This measure can be applied to communities of different sizes (not all of whose members need have a view on a given issue) and to questions whose answers have many different sorts of structure. Disagreement is weighted by a measure of 'distance' between any two answers, which makes alpha particularly helpful in comparing questions with different numbers of answers. With such a metric in hand, one can use a version of alpha to measure communal degree of agreement with a specific answer. For our purposes some rescaling may be useful (e.g. imposing a lower bound of zero and then squaring).

For the purposes above, what is needed is a variant on Krippendorff's alpha that measures communal degree of agreement with a specific answer. Some rescaling may be useful (e.g. imposing a lower bound of zero and then squaring).

But something like convergence to the truth is required in order that the convergence constitutes progress and not regress.

## 2. Argument for the central thesis

Here is an argument for the central thesis. It has two premises, an empirical premise and a bridging premise.

(1) Empirical premise: There has not been large collective convergence on the big questions of philosophy.
(2) Bridging premise: If there has not been large collective convergence on the big questions of philosophy, there has not been large collective convergence to the truth on the big questions of philosophy.
(3) Conclusion: There has not been large collective convergence to the truth on the big questions of philosophy.

The argument is valid. The bridging premise may look like a logical truth, but given how large collective convergence is defined above, it is not. The antecedent of this conditional premise says that there is less convergence on *dominant* answers to the big questions in philosophy than in the hard sciences, while the consequent says that there is less convergence on *correct* answers to those questions. These may come apart in a way that renders the conditional false if there has been strong convergence to the false (or strong convergence uncorrelated with truth) in the sciences along with weak convergence to the truth in philosophy. Still, given that convergence in science is largely convergence to the truth, or merely that convergence in science is at least as likely to be convergence to the truth as convergence in philosophy, then the premise is plausible.[6]

One may worry that because most scientific theories eventually turn out to be false, most convergence in science will be convergence to the false. This worry is less pressing if we formulate questions in terms of a small number of mutually exhaustive choices, as suggested earlier. This way, convergence to the truth will require only convergence on a correct coarse-grained class of theories, rendering it much more plausible that much convergence in science has been convergence to the truth. Alternatively, if questions allow an open-ended range of answers, then invoking a metric for distance between

[6] Thanks to Hedda Hassel Mørch and Rory Madden for pointing out ways in which the bridging premise could turn out to be false.



https://doi.org/10.1017/S0031819114000436 Published online by Cambridge University Press

answers (also suggested earlier) will allow that agreement on strictly speaking false theory that is relatively close to the truth constitutes a sort of convergence to the truth.

The main work in the argument is done by the empirical premise. I take it that it will be plausible to those with passing familiarity with philosophical and scientific practice. Still, it makes a sociological claim and cannot be decisively settled from the armchair. We do not have all the empirical data required for a systematic investigation of the premise, but we have some of it.

The 2009 PhilPapers Survey[7] surveyed professional philosophers on answers to thirty important questions in philosophy. The survey was sent to the members of 99 leading departments of philosophy (largely specializing in analytic/Anglocentric philosophy) in North America, Europe, and Australasia. About 47% of the 2000 or so recipients of the survey filled out and returned the survey. Questions were posed as a choice between two, three, or four options. Respondents could indicate that they 'accept' or 'lean toward' one option, or give a variety of 'other' answers (e.g., unfamiliar with the issue, the question is too ambiguous to answer, there is no fact of the matter, accept another option, and so on). The results (collapsing 'accept' and 'lean toward' answers, and collapsing 'other' answers) were as follows.

(1) A priori knowledge: yes 71%, no 18%, other 11%.
(2) Abstract objects: Platonism 39%, nominalism 38%, other 23%.
(3) Aesthetic value: objective 41%, subjective 35%, other 24%.
(4) Analytic/synthetic distinction: yes 65%, no 27%, other 8%.
(5) Epistemic justification: externalism 43%, internalism 26%, other 31%.
(6) External world: non-skeptical realism 82%, skepticism 5%, idealism 4%, other 9%.
(7) Free will: compatibilism 59%, libertarianism 14%, no free will 12%, other 15%.
(8) God: atheism 73%, theism 15%, other 13%.
(9) Knowledge claims: contextualism 40%, invariantism 31%, relativism 3%, other 26%.
(10) Knowledge: empiricism 35%, rationalism 28%, other 37%.
(11) Laws of nature: non-Humean 57%, Humean 25%, other 18%.

[7] Bourget, D. and Chalmers, D. J. 'What do philosophers believe?', *Philosophical Studies* (2013, Online First).

(12) Logic: classical 52%, non-classical 15%, other 33%.
(13) Mental content: externalism 51%, internalism 20%, other 29%.
(14) Meta-ethics: moral realism 56%, moral anti-realism 28%, other 16%.
(15) Metaphilosophy: naturalism 50%, non-naturalism 26%, other 24%.
(16) Mind: physicalism 57%, non-physicalism 27%, other 16%.
(17) Moral judgment: cognitivism 66%, non-cognitivism 17%, other 17%.
(18) Moral motivation: internalism 35%, externalism 30%, other 35%.
(19) Newcomb's problem: two boxes 31%, one box 21%, other 47%.
(20) Normative ethics: deontology 26%, consequentialism 24%, virtue ethics 18%, other 32%.
(21) Perceptual experience: representationalism 32%, qualia theory 12%, disjunctivism 11%, sense-datum theory 3%, other 42%.
(22) Personal identity: psychological view 34%, biological view 17%, further-fact view 12%, other 37%.
(23) Politics: egalitarianism 35%, communitarianism 14%, libertarianism 10%, other 41%.
(24) Proper names: Millian 34%, Fregean 29%, other 37%.
(25) Science: scientific realism 75%, scientific anti-realism 12%, other 13%.
(26) Teletransporter: survival 36%, death 31%, other 33%.
(27) Time: B-theory 26%, A-theory 16%, other 58%.
(28) Trolley problem: switch 68%, don't switch 8%, other 24%.
(29) Truth: correspondence 51%, deflationary 25%, epistemic 7%, other 17%.
(30) Zombies: conceivable but not metaphysically possible 36%, metaphysically possible 23%, inconceivable 16%, other 25%.

The degree of the disagreement here is striking, if unsurprising. Only one view (non-skeptical realism about the external world) attracts over 80% support. Three views (a priori knowledge, atheism, scientific realism) attract over 70% support, with significant dissent, and three more views attract over 60% support. On the other 23 questions, the leading view has less than 60% support.

Admittedly, not all of the questions are among the 'big questions' of the past, but certainly some are: the questions about the external world, free will, god, knowledge, meta-ethics, metaphilosophy,

mind, and normative ethics, for example. Only two of these (external world, god) have views with over 60% support (and in the case of the external world question, the consensus is somewhat misleading, since arguably the biggest question is *how* we know about the external world).

For fuller data to adjudicate the central thesis, we would need the results of the PhilPapers Survey not just in 2009, but at regular intervals in the past: 1909, 1809, and so on. At each point we would need to ask members of the philosophical community first, what they take to be the big questions of philosophy, and second, what they take to be the answers to those questions as well as to big questions from past surveys. We would also need to have analogous longitudinal surveys in other fields: the MathPapers Survey, the PhysPapers survey, the ChemPapers Survey, the BioPapers Survey, and so on. And we would need a reasonable measure of agreement at a time. I predict that if we had such surveys and measures, we would find much less convergence on answers to the big questions suggested by past surveys of philosophers than we would find for corresponding answers in other fields.

Some partial data is given by the 23 problems that David Hilbert posed for mathematics in 1900.[8] Around ten of these 23 problems have been clearly solved, leading to universal consensus, and seven have been partially solved, leading to partial consensus. A similar pattern could reasonably be expected in physics, chemistry, and biology. We can compare these results to the problems in Bertrand Russell's 1912 *The Problems of Philosophy*.[9] None of these have led to universal convergence and almost none have led to anything close.

Of course one can object to the thesis in various ways. One could argue that there has been more convergence on the big philosophical questions of the past than these case studies suggest. Alternatively, one could argue that there has been less convergence on the big scientific questions of the past than they suggest.

A version of the first objection springs from the observation that disciplines such as physics were once considered part of philosophy. If we go back to a time before the split, then insofar as the big questions of physics are among the big questions of philosophy at that point, high convergence on the former will lead to significant

[8] Hilbert, D. 'Mathematical problems', *Bulletin of the American Mathematical Society* **8** (1902): 437–79; Yandell, B.H. *The Honors Class: Hilbert's Problems and their Solvers* (A.K. Peters, 2002).

[9] Thanks to Jeremy Goodman for suggesting the Hilbert/Russell comparison.

convergence on the latter. Still, insofar as physics was just a proper part of philosophy, and one more susceptible to convergence than the other parts, one would still expect convergence on the former to produce less convergence on the latter. It is also not entirely clear that philosophy *as we understand it* should get the credit for the convergence in physics: what was called philosophy in the past was arguably a different and broader field.

In any case, one can bypass this objection by focusing on a point after the split between physics and philosophy: 1809 or 1909, say. The objector might respond that now there will be less convergence in philosophy only because we have split off the parts of it that have made most progress. But this is to concede the central thesis and argue for a certain explanation of it, one that I consider later in the paper. Another response is that there were further splits after this point: psychology, logic, linguistics, and economics, for example. Still, I think that the questions resolved by these areas constitute a small enough fraction of the big questions of philosophy in 1809 or 1909 that even if philosophy gets credit for them, this will not bring the level of convergence in philosophy close to the corresponding level in the hard sciences.

Another version of the first objection suggests that some big philosophical questions of the past have reached consensus and so have dropped off the list of big questions even without their own disciplines branching off. Perhaps something like this is plausible for some moral and political questions, for example, such as the question of whether all people are equal, where convergence within philosophy reflects convergence in society more generally. And there may have been questions that were regarded as truly important in a period (about the viability of certain versions of idealism, say) on which there is now a consensus view. The existence of questions like this helps to make a case against a glass-empty thesis. But where a glass-half-empty thesis is concerned, it suffices to note that the proportion of questions like these is lower in philosophy than in the hard sciences.

As for the second objection, one could argue that many of the big questions of the hard sciences are themselves philosophical questions and have seen low convergence: questions about the interpretation of quantum mechanics, for example, or about the locus of natural selection. But as long as some of the big questions of the hard sciences are not philosophical questions, as is surely plausible, and as long as these are more susceptible to convergence, then we would still expect the central thesis to be true. Furthermore, insofar as this objection relies on a contrast between philosophical and nonphilosophical

questions, it tends to reinforce the underlying contrast in convergence.

It could also be suggested that numerous big nonphilosophical questions in the hard sciences have met with low convergence: questions about the origins of life, for example. This is surely right, but it remains plausible that enough have seen major convergence that there is still a significant difference between the nonphilosophical and the philosophical. The case of the Hilbert problems for mathematics brings this out. The problems are mostly nonphilosophical, and although some of them are unsolved, the overall convergence on them has been quite dramatic. Something similar plausibly applies in physics, chemistry, and biology.

A final objection is that even though there is more agreement now in the hard sciences than in philosophy, these sciences may also start from a position of more agreement, resulting in a smaller *increase* in agreement in the sciences than in philosophy. This hypothesis is mathematically consistent, but I do not think it is especially plausible. Especially given a measure of agreement that is biased toward universal agreement, as discussed earlier, and given that there is considerable disagreement over the big questions of a time at that time, then the various cases in which the hard sciences (unlike philosophy) approach universal consensus will tend to yield greater overall convergence as well.

## 3. The varieties of progress

Despite this lack of convergence, it is hard to deny that the insights of Plato and Aristotle, Hume and Kant, Frege and Russell, Kripke and Lewis have involved significant philosophical progress. Correspondingly, my glass-half-empty thesis is compatible with many different glass-half-full theses, asserting the existence of various forms of progress in philosophy. We can systematize various such theses by dropping the central requirements of my central thesis one at a time.

Drop 'large': There has been (non-large) collective convergence to the truth on big questions of philosophy. It is plausible that there has been major convergence on answers to a small number of the big questions of philosophy. In questions about god there appears to have been major convergence toward atheism, for example. It is also plausible that there has been minor convergence on answers to many other questions, such as toward physicalism about the mind. Of course whether one counts this convergence as convergence to

the truth will depend on one's own philosophical views. Theists and dualists will hold that the convergence constitutes regress rather than progress. But if we assume optimistically that the convergence is indeed convergence to the truth, it may be that 10–20% more philosophers have true beliefs about the answers to the big questions of 1809 in 2009 than in 1809. If so, that is a sort of progress. Still, it remains plausible that convergence is greater in other areas.

Drop 'collective': There has been large (non-collective) convergence to the truth on the big questions of philosophy. The central thesis is consistent with the claim that various individuals or subcommunities have themselves had large convergence to true answers on the big questions. For example, on my more optimistic days I can convince myself that over time I have converged to the truth on many of these questions. But if so, sadly, it has not led to collective convergence to the truth. Likewise, perhaps groups such as the logical empiricists or the Oxford realists have converged on the truth. But again, community-wide convergence has not ensued. Perhaps there has even been a large amount of community-wide convergence at certain local temporal periods, but if so, this convergence has not persisted over time.

Drop 'big': There has been large collective convergence to the truth on (non-big) questions of philosophy. There has been large convergence on various smaller theses: the thesis that knowledge is not justified true belief, for example, and the thesis that conditional probabilities are not probabilities of conditionals. As van Inwagen suggests in the passage above, we are especially good at converging on negative theses that rule out certain specific views. There is also often convergence on conditional theses, asserting conditional connections between views. But I take it that these are not really answers to the big questions of philosophy.

Drop 'convergence to the truth': There have been large collective advances (not involving convergence to the truth) on the big questions of philosophy. Certainly there are many forms of philosophical progress that do not involve convergence to the truth. It is plausible that we have a greatly increased understanding of the issues underlying the big questions. We better understand the reasons for accepting and rejecting key philosophical theses. We have come to explore new views and new areas of philosophical space that we had not even conceived of earlier. We have developed new methods and better arguments. In some cases we have applied philosophy to the world. These are all certainly forms of progress. I simply note that they have not been accompanied by large collective convergence to the truth.

I want to stress that I am not simply equating progress with convergence to the truth. I am a pluralist about progress: there are many values that can be realized through philosophy, and there are many ways of advancing and realizing those values. Attaining the truth is certainly not the only such value. Still, it is certainly one such value. It follows that progress toward the truth is one form of philosophical progress.

More strongly, I think a case can be made that attaining the truth is the primary aim at least of many parts of philosophy, such as analytic philosophy. After all, most philosophy, or at least most analytic philosophy, consists in putting forward theses as true and arguing for their truth. I suspect that for the majority of philosophers, the primary motivation in doing philosophy is to figure out the truth about the relevant subject areas: What is the relation between mind and body? What is the nature of reality and how can we know about it? Certainly this is the primary motivation in my own case. So I am sympathetic with the claim that progress toward the truth has a certain primacy among the forms of philosophical progress. But even if one denies this, it is hard to deny that it is among those forms.

It is sometimes said that an obsession with truth reflects an overly scientistic conception of philosophy. We should not think of philosophy as a quest for the answers. Instead it is a quest for something else: understanding, clarity, enlightenment. I agree that these are goals worth pursuing, and that philosophy can help us pursue them. And I can see why, in the absence of answers to philosophical puzzles, it might seem especially appealing to focus on these goals instead. Still, I think we should acknowledge that this reaction involves something of a lowering of our sights for philosophy. At least pretheoretically, many of us get into philosophy looking for truth and looking for answers. One can argue that this hope is naive: truth and knowledge are not to be had in philosophy, and one should settle for something different. But even if so, both the hope and its naivete are worth marking.

Why is *convergence* to the truth important, and why should we be concerned about its absence? One obvious answer is that we value knowledge, agreement is required for knowledge, and convergence goes along with increases in knowledge. A strong version of this view, suggested by van Inwagen's discussion, is that where there is sufficient disagreement among experts, no individuals can be said to know the truth. Even if some individuals have hit on good arguments for true conclusions, how can they have justified confidence that these are good arguments, when so many of their peers disagree? I am not so sure: I think that at least in some cases,

a good argument can ground an individual's knowledge of a conclusion even when peers reject it. For example, I think that the presence of any number of peers who deny the existence of consciousness would not undermine my knowledge that I am conscious. Likewise, it would not undermine arguments that take this claim as a premise.

But even if agreement is not required for individual knowledge, some degree of agreement is plausibly required for *collective* knowledge. If the community of experts on a question has serious disagreement over the answer to that question, then that community cannot be said to collectively know the answer to that question, and nor can the broader community of which they are a part. Even when some individuals know the answer to a question, this individual knowledge will not usually suffice for collective knowledge, except perhaps in special circumstances such as when the community defers to these individuals.

Furthermore, we value collective knowledge. One reason that progress of the hard sciences has been so impressive is that it has plausibly enabled us – the community of inquirers – to collectively know the answers to those questions. But in the absence of sufficient agreement on the answers to philosophical questions, we cannot be said to have collective knowledge of those answers.

Of course one can argue over just what degree and pattern of agreement is required for collective knowledge. But it is highly plausible that the kind of disagreement that we observe over the answers to the big questions of philosophy suffices to undermine any claims of collective knowledge of the answers to most of those questions. Perhaps one could argue that in the survey above, a few views (non-skeptical realism about the external world, atheism, a priori knowledge) display the sort of consensus that allows collective knowledge. But even that claim would be bold, and the extension to claims with less consensus (physicalism and compatibilism, say) seems so bold as to be implausible. So I take it that the difference in agreement on the big questions in science and philosophy reflects a significant difference in the collective knowledge that we have attained. Likewise, the difference in convergence on the big questions reflects a significant difference in the increase of collective knowledge over time.

This is not to deny that we have attained a great deal of collective knowledge in philosophy. As Timothy Williamson has said,[10] we knew much more in 2004 than in 1964, much more in 1964 than in

[10] Williamson, T. 'Must do better', in *Truth and Realism* (eds) P. Greenough & M. Lynch (Oxford University Press, 2006).



https://doi.org/10.1017/S0031819114000436 Published online by Cambridge University Press

1924, and so on. But this collective knowledge typically does not involve answers to the big questions. It is mainly knowledge of the answers to smaller questions, of negative and conditional theses, of frameworks available to answer questions, of connections between ideas, of the way that arguments bear for and against conclusions, and so on. In the absence of convergence on the big questions, collective knowledge of the answers to those questions eludes us.

## 4. Philosophical argument

I now turn to the central question: why isn't there more progress in philosophy? And in particular: why is there less convergence in philosophy than in the hard sciences?

An initial explanation, though perhaps this is merely an articulation of the phenomenon, lies in the relative power of the methods used in these domains. The hard sciences have methods – proof in the case of mathematics, and the observational/experimental method in physics, chemistry, and biology – that have the power to compel agreement on the answers to the big question. Philosophy has a method – the method of argument – that does not.

What is the difference between these methods? One difference in that the methods of experiment and proof start from widely agreed premises – observations in science, axioms in mathematics – and proceed from there to strong and surprising conclusions. We aspire to do this in philosophy too: witness Russell's remark that the point of philosophy is to start with something so simple as not to seem worth stating, and to end with something so paradoxical that no one will believe it.[11] But in practice, widely agreed premises rarely suffice to ground strong and surprising conclusions in philosophy.

There are certainly many arguments for strong conclusions in philosophy. But in the great majority of cases, they have premises that opponents can deny without too much cost, or inferences that opponents can reject without too much cost. (I focus mainly on premises, but everything I say also applies to inferences, for example by turning nondeductive inferences into tacit premises of deductive arguments.) Sometimes the denied premise is antecedently plausible, and the denial somewhat surprising. But even then the denial rarely

[11] Russell, B. 'The Philosophy of Logical Atomism', *The Monist* **28** (1918): 295–527 (Reprinted as *The Philosophy of Logical Atomism* (Taylor and Francis, 2009)).

has the implausibility of denying a mathematical axiom, or of denying a well-replicated experimental observation. So these denials are usually tenable, at least in a broadly sociological sense of tenability.

Let us say that *consensus* premises (and inferences) are those that are regarded by the community as undeniable, or at least as incurring enormous cost for anyone who denies them. A consensus premise might be denied by a few outliers, but it cannot be subject to widespread disagreement within the community. Let us say that an argument that uses only consensus premises and inferences is a decisive argument. (Note that consensus premises and decisive arguments are both defined in sociological terms.) Then the claim is that while there are decisive arguments for strong conclusions in the sciences, there are relatively few such arguments in philosophy.

There are certainly some consensus premises in philosophical arguments. After all, these premises can include mathematical axioms and empirical observations themselves, as well as the theorems and theories that are grounded in them. But mathematical premises and empirical observations alone almost never suffice to draw strong philosophical conclusions. Further premises or inferences are required to bridge from science and mathematics to philosophy, and these premises and inferences are typically deniable.

In addition to mathematical axioms and empirical observations, there are some philosophical intuitions that are extremely difficult to deny. But these intuitions are not so common (many antecedently plausible intuitions turn out to be deniable), and where they exist, it is typically difficult to draw strong philosophical conclusions from them. There are some cases where these intuitions, perhaps in conjunction with mathematical and empirical claims, allow us to draw strong and surprising conclusions. This works particularly well for negative theses, where intuitions and formal models can generate counterexamples to positive theses or other reasons to reject them. Gettier's argument from an intuition about a case to the conclusion that knowledge is not justified true belief is one example. Lewis's formal argument that conditional probabilities are not probabilities of conditionals is perhaps another. But it is notable that these negative conclusions fall far short of answers to the big questions of philosophy. Almost any argument for a positive answer to these questions involves deniable premises.[12]

[12] To gather data here, I ran an informal Internet survey of philosophers, asking for arguments that are near-universally regarded by philosophers as establishing their conclusions. Further candidates included the forcible-organ donation argument against simple versions of utilitarianism, Kripke's



https://doi.org/10.1017/S0031819114000436 Published online by Cambridge University Press

For most practitioners of philosophy, the phenomenon of premise deniability is familiar from both sides. As the old saying goes: one person's modus ponens is another person's modus tollens. When we give arguments for our views, we are frustrated to find opponents biting the bullet by rejecting what we took to be a plausible premise, without this serving as any sign of defeat. When we address arguments against our views, we sometimes work backwards from our rejection of the conclusion to see which premises we have to deny, and we deny them. In the best cases, we learn something from this, and we take on commitments that we might have antecedently found surprising. But these commitments are rarely untenable to maintain.

As a result, philosophical arguments typically lead not to agreement but to sophisticated disagreement. Advocates of a view learn what extra commitments they need to take on to avoid the arguments. Bad versions of a view are rejected and sophisticated versions are developed in their place. This leads to a sort of eliminative progress where areas of philosophical space are eliminated, but only in small fragments at a time. It is rare for a major general view (materialism or dualism, compatibilism or incompatibilism, utilitarianism or deontology) to be eliminated in this way. Instead, there are large

---

argument that necessity comes apart from apriority, Gödel's argument against versions of mathematical formalism, the argument from evil against theism, the model-theoretic argument against global descriptivism, the perfect actor argument against logical behaviorism, the multiple-realizability argument against the identity theory, Goodman's argument against purely formal inductive logic, arguments from relativity against presentism, Frankfurt's argument that moral responsibility does not require the ability to do otherwise, Hart's argument against Austin's command theory of laws, Russell's refutation of Frege's Basic Law V, Moore's open question argument against analytic naturalism, Putnam's argument for externalism about meaning, Descartes' cogito, and many others.

It is striking that the great majority of these arguments are naturally regarded as arguments for negative conclusions, in that they are arguments against fairly specific views. Of course the negative/positive distinction is not entirely clear, but we have a reasonable intuitive grasp on it. A few conclusions have a positive flavor: one's existence (the cogito), externalism (Putnam), and perhaps the necessary a posteriori (Kripke) and atheism (the argument from evil). But the first three are at best marginal candidates for answers to big questions, and the survey data suggests that the second and fourth are at best marginal cases of near-universal agreement. All this reinforces the point that decisive arguments in philosophy are rare, that decisive arguments for positive views are even rarer, and decisive arguments for positive answers to the big questions are so rare as to be almost nonexistent.

surviving fragments containing the views needed to avoid the arguments (type-B materialism with the phenomenal concept strategy, source incompatibilism, two-level utilitarianism, theism without unrestricted benevolence or omnipotence). The same sort of elimination, fragmentation, and refinement often recurs at these lower levels. The views that survive yield a sort of fractal structure in philosophical space, akin to the Mandelbrot set with its intricate complexities at all levels, but in which large regions of space are rarely eliminated entirely.

This phenomenon might strike one as a philosophical analog of the Duhem-Quine thesis, in a version saying that any scientific theory can be made compatible with any evidence by appropriate adjustment to the background assumptions that bridge between theory and evidence. But in practice, scientific theories are often decisively rejected in light of evidence, with revised consistent versions of the theories being rejected as untenable. In effect, some (nondeductive) inferences from evidence to scientific theory have consensus status. Theories are ruled out not by consensus evidence alone, but by consensus evidence plus consensus inferences. In the philosophical case, however, consensus evidence plus consensus inferences are much less powerful. When someone argues against a philosophical theory, there is usually at least a revision of the theory that is not just consistent but tenable in light of the consensus evidence. Those who argue against a philosophical view sometimes accuse their resourceful opponents of holding onto a degenerating research program, but it is typically much harder to make this charge stick in philosophy than in science. This could be because philosophers apply laxer standards to their theories, so that inferences that have consensus status among scientists do not have it among philosophers, but more plausibly it is because inferences that meet scientific standards provide weaker reasons to accept philosophical conclusions than to accept scientific conclusions.

It might also be objected that in science, positive theories are not usually established by single experiments, but by many experiments collectively. By parity, we might hope that even if positive philosophical theories are not established by single arguments, they might be established by a number of arguments collectively. There are perhaps a few cases of negative theses being established this way: the rejection of sense-datum theories of perception may be an example. But even these cases are rare, and positive cases are even rarer. In practice, if an opponent can reject individual arguments for a thesis without too much cost, they can usually reject collections of arguments without much cost too.

Does this mean that all philosophical arguments for positive theses are unsuccessful, as van Inwagen has suggested?[13] (Van Inwagen talks about substantive theses, but his discussion suggest that these are required to be positive theses.) This depends on what one means by success. If one defines success in sociological terms, so that success requires convincing almost everyone in a community, then we have seen that at best very few philosophical arguments for positive theses have been successful in our community. Van Inwagen defines success in idealized epistemological terms: a successful argument for a proposition $p$ is one that would convince an audience of ideal reasoners who are initially agnostic concerning $p$, in the presence of an ideal opponent of $p$. I do not think that the sociological observations above (or the sociological observations that van Inwagen appeals to) come close to establishing that no philosophical arguments are successful in that sense. (See McGrath and Kelly forthcoming for more on this theme.) Human beings are simply too far from ideal for that conclusion to follow.

It also does not follow from anything I have said that all philosophical arguments are question-begging, or that they are dialectically powerless. Even when arguments have deniable premises, they often have dialectical power, in that their premises have antecedent support that does not rest on considerations too close to the conclusion. In such a case the argument does not beg the question. Even though a sophisticated and committed opponent will deny the premise, the argument might well move an agnostic observer to accept the conclusion. In practice, we often use this sort of dialectical power as a criterion for a good argument that many philosophers can agree on, even if they disagree on the argument's ultimate persuasiveness.

I am also not saying that these arguments cannot produce knowledge. Deniable premises may nevertheless be known by many people to be true. As before, while too much disagreement over a claim may undermine collective knowledge of that claim, it need not undermine individual knowledge of that claim. Likewise, an argument can ground individual knowledge even when peers reject it. This applies all the more where non-peers are concerned. For all I have said, some arguments may have premises and inferences that can only be denied unreasonably, or by nonideal reasoners. If so, these arguments may well produce knowledge in beings more reasonable than the deniers.

[13] Van Inwagen, P. *The Problem of Evil: The Gifford Lectures* (Oxford University Press, 2006).

So it is not straightforward to draw conclusions about lack of normative force from premises about lack of sociological success. There is perhaps an intermediate normative notion, defining a successful argument as one with the power to persuade all *competent* agnostics, where competence is some reasonably high but nonideal standard of rationality that many human philosophers meet. There is good reason to think that few philosophical arguments for positive conclusions persuade all competent philosophers, or even all competent agnostics. On the face of it, disagreement on big questions among the most able philosophers (by any reasonably neutral measure) is about as rife as disagreement among philosophers more generally. This suggests that most philosophical arguments are not successful in the normative sense tied to competence, even if they are successful in the other normative senses.

Burton Dreben once memorably said to me (on the only occasion that I met him, in St. Louis around 1994): 'Great philosophers don't argue'. He went on to elaborate that none of Frege, Russell, Wittgenstein, Carnap, or Quine really give arguments for their views. Of course this is not strictly true, but I think his point was that for these philosophers, the real work is not done by arguments for a thesis, but by the thesis itself, or the framework it is embedded in. A refined version of his claim (suggested to me by Gene Callahan) might say: great philosophers may argue, but their arguments are not what makes them great. A part of Dreben's thought, as I understood it, was that since arguments are so easily rebutted, giving arguments is a sign of weakness. It's better to simply assert and develop a thesis. Then one's readers have to engage with the thesis itself, without the cheap distraction of rebutting arguments for the thesis.

(Rawls elaborates on Dreben's views in a somewhat different direction: 'Burt would not, of course, deny the plain fact that philosophers make many complicated arguments. But he thinks that at bottom there are no arguments one philosopher can use to convince another of a metaphysical point. At the basic level, philosophers simply rely on and appeal to different "data". It is a standoff with no resolution by argument. Burt has said that Quine is a metaphysician, a metaphysician of science. By that he means that Quine doesn't argue for physicalism, or scientific realism. He assumes it and works out his view from there.'[14])

[14] Rawls, J. 'Afterword' in *Future Pasts: The Analytic Tradition in Twentieth-Century Philosophy*, (eds) J. Floyd & S. Shieh (Oxford University Press, 2001).

I have found it impossible to follow Dreben's advice myself. In my work I am a compulsive arguer, which no doubt leaves me subject to a *modus tollens* from Dreben's thesis. But certainly it is rare that these arguments bring a large sector of the population around. This is especially so when many of the people already have firm commitments, as on issues such as the mind–body problem and the theory of meaning: here it is hard to do more than bringing around a few people here and there. On issues where people are initially agnostic or their commitments are weak, there can be more movement.[15] Even here the fact that such movement is limited reinforces the basic point.

The upshot is that consensus in philosophy is as hard to obtain as it ever was, and decisive arguments are as rare as they ever were. To me, this is the largest disappointment in the practice of philosophy. Once one has been doing philosophy for a while, one no longer expects arguments to produce agreement, and one deems an argument good when it merely has some dialectical power. But this is an adjustment of expectations in response to a disappointing reality. Antecedently to doing philosophy, one might have hoped for something more.

## 5. New philosophical methods

Faced with the failure of traditional philosophical methods, we might look to new methods. Occasionally, new methods developed by thinkers who considered themselves philosophers have helped to resolved questions once considered philosophical: witness the development of logic, physics, psychology, and so on. It is natural to hope that new methods might produce further progress.

In the last century or so, many new philosophical methods have been developed and many old methods have been refined, in

[15] To indulge in autobiography: I have the sense that my arguments with Andy Clark for the extended mind thesis (an area where prior commitments were relatively weak) may have brought more people around than my arguments against physicalism or for two-dimensional semantics (areas where prior commitments are strong). Even there I suspect that the thesis and the framework have brought around as many people as the arguments. Perhaps most effective of all has been the argument in 'The Matrix as Metaphysics', which brings many people around to the view that if we are in a matrix scenario or that if we are brains in vats, most of our beliefs are true. (At least it does this in lecture presentations and informal discussions; there has been relatively little discussion of the argument in print.) Although people find this view initially counterintuitive, it turns out that their antecedent commitment was weak.

order to help reach philosophical conclusions. Empirical philosophy draws on empirical science. Formal philosophy draws on formal reasoning. Linguistic philosophy draws on the analysis of language. Phenomenology draws on phenomenological reflection. Feminist philosophy draws on analysis of the role of gender. Cross-cultural philosophy draws on multiple cultural traditions in philosophy. Experimental philosophy draws on the empirical study of philosophical judgments.[16]

All of these methods have led to new insights and to philosophical progress. All have led to new arguments for interesting conclusions. But manifestly, none of these methods have led to recent convergence on answers to the big questions of philosophy. In the wake of each of these methods, philosophical disagreement is as rife as it ever was. Even within a tradition, there are few cases where the big questions are regarded as settled. Instead, these methods have led us again to more sophisticated versions of old disagreements.

In many cases, the basic problem is that of premise deniability. In the case of empirical and formal philosophy, we have already seen that empirical and formal results must be combined with further bridging premises to settle a philosophical question. In most cases, it turns out that these bridging premises can be denied. Often they are about as controversial as the conclusions they aim to establish. In some cases, empirical and formal results help to settle relatively small questions, as well as introducing and addressing new important questions. But when they are brought to bear on the big questions, it is rare that they do much to produce consensus.

There are some partial exceptions: perhaps the bearing of evolution on theism, the bearing of relativity on presentism, and the bearing of Gödel's theorem on mathematical formalism. But these exceptions are not especially common, and even in these cases, there are modified versions of the relevant views that have retained numerous serious adherents. What exceptions there are seem mainly to fall into two classes. First, there are cases where empirical methods bear strongly on areas of philosophy that focus on concrete reality, such as subfields of metaphysics and the philosophy of science. Second, there are cases where formal methods bear strongly on areas that deal with formal questions, such as the philosophy of mathematics and logic. Of course both sorts of method are often brought to bear on other

[16] Then there are many other methods that I am not competent to discuss. For example, Nielsen ('Can There Be Progress in Philosophy?' *Metaphilosophy* **18** (1987): 1–30) suggests that the one hope for progress in philosophy is critical theory.

areas – normative areas such as ethics and epistemology, for example – but cases in which they produce consensus are much rarer.

One might think that the philosophy of mind would be an exception, given the obvious connections to neuroscience and psychology. But even here, these sciences seem to have left the big questions – the problems of consciousness and intentionality, of mental causation and free will – wide open. Certainly there have been arguments from neuroscience and psychology to views about these problems, but in most cases the bridging premises required have been as controversial as most other philosophical claims. What has resulted is a greatly increased sophistication with scientifically informed versions of the relevant views, but not much more in the way of consensus. Perhaps the greatest bearing on these big questions about the mind has come not from these fields but from physics, where the evidence for causal closure at the microphysical level has put serious pressure on views such as interactionist dualism. But even here many have resisted the pressure, and in any case the denial of interactionism does not really amount to a positive view.

Other new methods do not even offer premises with the relative security of empirical and formal premises. In phenomenology, for example, the key phenomenological premises are typically as deniable as any other philosophical premises. Something similar may apply to feminist and cross-cultural philosophy, while other methods may exhibit a mix of the two patterns above.[17] Some methods, such as feminist philosophy and experimental philosophy, have played a crucial critical role, but the upshot has been largely to lessen our confidence about the answers to the big questions rather than to strengthen it.

Of course, new methods are always being developed. It still happens that issues gradually migrate from philosophy to science as methods become more rigorous and decisive: two recent examples include the development of formal semantics and the ongoing development of a science of consciousness. Still, even in these cases it would be hard to say that the new methods have led to consensus on the biggest philosophical questions that preceded their development. So while we can hope for further methods that produce convergence on the big questions, these methods will have to go well beyond what we have seen over the last century or so.

[17] For my take on the power and limits of experimental and linguistic philosophy, see http://consc.net/papers/xphi.pdf and http://consc.net/papers/langphil.pdf respectively.

## 6. Explanations

So far I have given a very partial explanation of the relative lack of convergence in philosophy. There is less convergence in philosophy because the philosophical method has less power to compel agreement, and it has less power because of the phenomenon of premise deniability: arguments for strong conclusions in philosophy (unlike science and mathematics) almost always have premises or inferences that can be rejected without too much cost.

Still, this explanation stays fairly close to the surface of the phenomenon. It is natural to ask for a deeper explanation. Why are arguments from consensus premises relatively powerless to settle the big questions of philosophy? And more generally, why is there so little convergence in philosophy?

(1) *Disciplinary speciation*. The most popular answer to this question, at least among philosophers, is that the field is subject to a sort of disciplinary speciation. As I have discussed already, many new disciplines have sprung forth from philosophy over the years: physics, psychology, logic, linguistics, economics, and so on. In each case, these fields have sprung forth as tools have been developed to address questions more precisely and more decisively. The key thesis is that when we develop methods for conclusively answering philosophical questions, those methods come to constitute a new field and the questions are no longer deemed philosophical. So it is only to be expected that the questions that remain are subject to less agreement than those in other disciplines.

There is certainly something to this explanation. The key thesis is plausible and the central thesis seems to follow from it. Still, I think there are some limits on this explanation.

I have already noted one limit: the fields that have split off have not always answered the big philosophical questions that preceded them. Psychology has not done much to settle the mind–body problem, for example, and linguistics has not really settled the deepest philosophical questions about meaning. Logic and physics have come closer, but even here it is arguable that they have not settled some of the biggest antecedent philosophical questions. Now, it might be said that the smaller philosophical questions that these fields settle nevertheless correspond to the big questions in the new fields, thereby explaining the central thesis about relative convergence. Still, one wants an explanation of why the *antecedent* big questions in philosophy have been so hard to answer. Insofar as these big questions have not been resolved by disciplinary speciation, then speciation cannot answer that question.



https://doi.org/10.1017/S0031819114000436 Published online by Cambridge University Press

A more general objection is that although the speciation thesis may explain, de dicto, why there is less convergence on big questions in philosophy than on big questions elsewhere, we also want a de re explanation, concerning those big questions, of why *they* in particular receive so little convergence. One hypothesis is that this is simply a matter of luck: all questions are equally apt for convergence, and through random luck some have received convergence (and thereby speciated) sooner than others. But setting aside this implausible hypothesis, the relative lack of convergence on these questions is presumably explained by something distinctive about those questions and their relation to us. We can then ask just what distinctive feature or features of these questions explains the lack of convergence. Speciation does not have the power to answer this question, so a further answer is required.

(2) *Anti-realism*. One answer is that there is no convergence to the truth because there are no objective truths to be had in the relevant domains. Where there is objective truth, it serves as a sort of magnet for convergence, but in its absence there is simply an unruly body of opinion which we should not expect to converge. Many philosophers will have sympathy to this line in some areas. I have sympathies with anti-realism about ethics and some questions in ontology. Still, this leaves plenty to be realist about. And even accepting moral anti-realism, say, leaves open why there is so little convergence on the question of moral realism itself. Of course there remains the possibility of global philosophical anti-realism, but this is not an especially plausible or attractive view.

(3) *Verbal disputes*. Another answer is that there is little convergence because participants are talking past each other. Each side is using key terms in different ways and each is correct where their own use of the term is concerned. In 'Verbal Disputes'[18] I argued that verbal disputes are common in philosophy. For example, I think many debates in the philosophy of free will and the philosophy of language have a significant verbal element. And I think that resolving verbal disputes can lead to philosophical progress. Still, often when we clarify the key terms in a partly verbal dispute, we find that a substantive dispute remains. And there is a core of fundamental questions (including many normative questions, as well the mind – body problem and other issues involving 'bedrock' philosophical concepts, in the terms of 'Verbal Disputes') for which the diagnosis of a verbal dispute seems quite implausible.

[18] Chalmers, D. J. 'Verbal disputes' in *Philosophical Review* **120** (2011): 515–66.

(4) *Greater distance from data*. An answer naturally suggested by the discussion of decisive arguments is that there is less convergence in philosophy than in science because philosophy tends to concern domains that are remote from clear data. To put this in a Quinean mode, philosophical theses are a long way from the periphery in the network of belief. Still, on the face of it, the same goes for many highly theoretical claims in science, for example concerning the distant past and the very small. And plausibly the same goes for mathematics. In that case one might point to mathematical axioms and intuitions as data, but this then raises the question of why we don't have analogous philosophical data to settle philosophical questions. So this option tends to relabel the problem rather than to solve it.[19]

(5) *Sociological explanations*. It is natural to suppose that sociological factors play a role in preventing convergence. When our arguments are not universally accepted, we often chalk this up partly to our opponents' professional background, or to false assumptions that are widespread in the profession, or to professional or emotional attachment to alternative views.

I think there is no denying that sociological factors play a major role in determining which philosophical views are widely accepted at a time. The unpopularity of the analytic-synthetic distinction in the decades after 'Two Dogmas of Empiricism' was certainly not uncorrelated with Quine's position of power in the profession and his impact on graduate students. Still, many of the relevant sociological factors are also at play in the sciences. So to explain a difference with the sciences, one has to either point to relevant sociological differences, or combine the sociological explanation with other distinctive features of philosophical questions.

Some potential sociological differences include the hypothesis that philosophers are rewarded for disagreement more than in the sciences, that they are more tolerant of dissent, or that they have been trained to have higher standards for acceptance of views. One could also point to differences in funding, training, and research

[19] It is also worth noting (as Larry Solum suggested to me) that the social sciences have much less convergence than the hard sciences despite being less remote from data than philosophy. An interesting general question is whether the lack of convergence in social sciences and in philosophy should receive different explanations or a uniform explanation. My suspicion is the former: for example, the complexity and messiness of social systems seems especially relevant in the social sciences but less relevant in philosophy.

structures. Still, it is hard to believe that the difference in convergence between, say, the human genome project and the mind – body problem merely comes down to these sociological factors. So sociological explanations work best when they are combined with further theses about the distinctiveness of philosophical questions. For example, one could suggest that the greater distance between data and philosophical theses makes it easier for sociologically grounded resistance to a thesis to get a grip.

(6) *Psychological explanations*: Closely related are psychological explanations, holding that there is something distinctive about human minds or about philosophers' minds that prevents convergence on philosophical questions. Perhaps there is some psychological flaw that prevents us from recognizing philosophical truth, for example. At some level some explanation like this must be part of the story: at least if ideal reasoners could converge on the truth, then our failure can be marked down to the nonideality of our reasoning. But now the crucial questions will be: what are the respects in which our reasoning is nonideal, and what are the respects in which philosophical questions are distinctive, such that this nonideality of our reasoning prevents us from converging to the truth on philosophical questions?

(7) *Evolutionary explanations*: It is sometimes suggested[20] that there is a Darwinian explanation for the lack of progress in philosophy. The rough idea is that we did not evolve to be good at philosophy, since in our evolutionary environment there were no selection pressures that favored philosophical ability or anything that strongly correlates with it. Perhaps there is something to this, though it would take some work to explain why the same does not apply to the ability to do abstract mathematics or highly theoretical science. In any case this sort of explanation will work best in conjunction with a psychological explanation, and raises the same crucial questions discussed under that topic.

I think all seven of these explanations may be partially correct. I do not think that they collectively provide a full explanation of the phenomenon as they stand, though. To do that, many of the details would need to be fleshed out. In particular, we still need a good account just what is distinctive about philosophical questions such that they lead to lack of convergence. Only (2) and (4) really address this, but (global anti-realism aside) (2) only applies in some cases, while (4) is too close to a restatement of the phenomenon. It ought to be possible to provide an account of this distinctiveness

[20] McGinn, C. *Problems in Philosophy* (Oxford University Press, 1993).

that meshes with psychological, evolutionary, and perhaps sociological explanations to provide a full account of the lack of convergence. But for now I think this remains an open question.

To sum up, it appears that what is needed is an account of how philosophical questions are distinctive, and an account of why questions of that sort produce little convergence. The most obvious version of the second account would be an account of relevant aspects of our psychology or sociology. That would yield a sort of two-component 'lock-and-key' explanation, suggesting an imperfect match between our key (human psychology or sociology) and an intellectual lock (philosophical questions). Alternatively, it one could perhaps give a version of the second account on which philosophical questions turn out to be 'objectively' hard and not just hard for humans. That would be a sort of lock-and-key account on which the key is characterized in terms of species-neutral epistemology rather than human psychology. It is not clear just what this objective hardness would come to, and it is not clear why we should expect that philosophical questions are objectively hard in this sense, so perhaps the human-centered version is more plausible. Either way, the lock and key strategy seems to be a promising one. We just need to give the right accounts of the lock and the key.

## 7. The prospects for further progress

Finally: what are the prospects for further philosophical progress? Is it possible that we may eventually converge to the truth on the big questions of philosophy?

To get a grip on this, we need to address the question of whether the answers to these questions are even knowable in principle, by sufficiently ideal reasoners. Here I will just flag my own positive view on this question. In *Constructing the World*, I argued for a scrutability thesis (called Fundamental Scrutability in the book) holding that all truths are a priori entailed by fundamental empirical truths concerning fundamental natural properties and laws. It follows (roughly) that if someone could know all the fundamental empirical truths and reason ideally, they could know all the truths, including all the philosophical truths.

Of course the scrutability thesis can be denied. If it is false, then even ideal reasoning from fundamental empirical truths may not enable us to know the philosophical truth. One could preserve a modified version of the thesis by expanding the fundamental truths in the base to include certain philosophical truths: fundamental



https://doi.org/10.1017/S0031819114000436 Published online by Cambridge University Press

normative and ontological principles, say. But then the fundamental truths themselves may lie beyond an ideal epistemological grasp. Either way, if philosophical truths are not scrutable from an appropriate basis, we should not expect convergence to the truth even in a community of ideal reasoners. This would be an extreme version of the view discussed above on which philosophical questions are objectively hard.

If the scrutability thesis is true, on the other hand, a more optimistic view ensues. The thesis does not entail that we can know all the philosophical truths, but it provides a useful way to classify the cases where we fall short, and more generally to classify cases where we fail to converge. First, there are cases of anti-realism about a domain, where there is no philosophical truth to know. Second, there are cases where multiple parties all know philosophical truths, but where verbal disputes get in the way of their recognizing their agreements. Third, there are cases where we are ignorant of relevant fundamental empirical truths. Fourth, there are cases where our reasoning is nonideal.

I think that many of the hardest cases in philosophy fall into the last category: questions whose answers are knowable by ideal reasoners, but not (yet) known by us. This then raises the key question: are the answers knowable or unknowable by humans?

McGinn[21] and van Inwagen[22] have advocated unknowability: humans are just not smart enough to answer the big questions. The idea is that there is some level of intelligence or aptitude that would suffice to answer these questions, but that humans fall below that level.

Van Inwagen argues for this conclusion as follows. He suggests that it is implausible that we are much above that level, given the lack of progress to date, and that it is antecedently improbable that we should be just barely at that level. So it is much more likely that the level lies above us. I am not so sure about this argument. I think we already know that for a vast range of questions, humans are just barely at the level for doing them well: scientific and mathematical questions, for example. Because of this, it is arguable that we lie at a special intelligence threshold at which an extraordinarily wide range of questions come to be within our grasp over time. It is not obvious whether or not philosophical questions fall within that range, but it is not obviously more likely that they do not than they they do.

[21] McGinn, C. *Problems in Philosophy* (Oxford University Press, 1993).
[22] Van Inwagen, P. *Metaphysics*, Third Edition (Westview Press, 2009).

If McGinn and van Inwagen are right, it remains open that we could answer philosophical questions by first improving our intelligence level, perhaps by cognitive enhancement or extension. Alternatively, we could construct artificial beings more intelligent than us, who will then be able to construct artificial beings more intelligent than them, and so on. The resulting intelligence explosion might lead to creatures who could finally answer the big philosophical questions.

If McGinn and van Inwagen are wrong, on the other hand, then we may eventually answer philosophical questions without radical cognitive enhancement. We may need to develop new methods, increased discipline, new sorts of insights, and perhaps there will need to be a conceptual revolution or two, but none of this will lie outside human capacity. It may turn out that there is a curve of increasing philosophical sophistication such that past a certain point on the curve, major progress is possible. We are not there yet, but we are working our way toward it.

It is not obvious whether McGinn and van Inwagen are right or wrong. The question of whether the big philosophical questions are humanly solvable is itself a big metaphilosophical question. Like other big questions in philosophy, it is one we do not currently know the answer to. Both answers to this metaphilosophical question seem to be open, and we do not currently have strong reasons to favor either one.

If we don't know which of these two options obtains, then I think to do philosophy we can make the working assumption that it is the second: the questions are answerable by us but as yet unsolved. Then we can simply do philosophy as well as we can, doing our best to come up with those new insights, methods, and concepts that might finally lead to us answering the questions. After all, we are still learning to do philosophy well. To see how far it can take us, we have to keep doing philosophy.

DAVID CHALMERS (chalmers@nyu.edu) *is a philosopher at New York University and Australian National University. He is author of* The Conscious Mind, The Character of Consciousness, *and* Constructing the World.



https://doi.org/10.1017/S0031819114000436 Published online by Cambridge University Press