Are There Any Successful Philosophical Arguments?

Sarah McGrath
Princeton University

Thomas Kelly
Princeton University

1. Introduction: a pessimistic verdict

There is no shortage of bad arguments in philosophy. On this point, even philosophers who agree about little else will find common ground. Indeed, given the traditional conception of philosophy as an argument-driven discipline, the frequency with which philosophers have produced and consumed bad arguments can seem like a depressing fact about the subject. Still, one might wonder: just how bleak is the situation?

If Peter van Inwagen (2006: 52) is correct, the situation is bleak indeed:

There are certainly successful arguments, both in everyday life and in the sciences. But are there any successful philosophical arguments? I know of none. (That is, I know of none for any substantive philosophical thesis.)

Although couched as a disavowal of knowledge, it is clear from the context that what van Inwagen says here is not intended as a confession of idiosyncratic ignorance on his part. Rather, it is offered as an expression of skepticism about whether there are any arguments of the relevant kind. Let’s call this

**The Pessimistic Verdict:** There are no successful philosophical arguments for substantive philosophical theses.¹

¹ It is clear enough that van Inwagen believes that The Pessimistic Verdict is true. For example, at (2006: 53) he reports his belief that “no philosophical argument for a substantive conclusion has the power to convert every member of an ideal and initially neutral audience to its
In this paper, we offer some critical reflections on van Inwagen’s case for The Pessimistic Verdict and his stimulating discussion of related issues, as presented in the third of his 2003 Gifford Lectures, “Philosophical Failure”.\(^2\) Central to van Inwagen’s case for The Pessimistic Verdict is his development and defense of a ‘criterion of philosophical success’: that is, an account of the conditions that a philosophical argument must satisfy in order to qualify as a successful argument. With this account in hand, he proceeds to argue that, plausibly, no philosophical argument for a substantive conclusion meets the relevant standard. We will proceed as follows. In Section 2, we raise some preliminary questions about The Pessimistic Verdict, focusing in particular on how the crucial notion of a \textit{substantive} philosophical thesis should be understood. In Section 3, we explore van Inwagen’s criterion of philosophical success and develop an objection to that account. In Section 4, we argue that even if we are wrong and van Inwagen’s account of philosophical success is correct, his case for The Pessimistic Verdict is not compelling. We close with some further remarks in Section 5.

A certain skepticism about the potency of philosophical argument is a recurring theme in van Inwagen’s writings (see, e.g., 2004: 334-340; 1990: 115). We have chosen to focus on “Philosophical Failure” because it represents his most sustained discussion of the issue. Because it will play some role in what follows, a few remarks about the context of that discussion are in order. The overarching project of van Inwagen’s Gifford Lectures is to show that “the argument from evil”—roughly, the argument that purports conclusion”. And as we will see in Section 3 below, van Inwagen holds that an argument is a success if and only if it satisfies this condition.

\(^2\)Ultimately published as \textit{The Problem of Evil} (Oxford University Press, 2006). In what follows, any page references refer to this published text, unless otherwise noted.
to establish the nonexistence of God on the basis of the existence of evil—is a failure. In this context, “Philosophical Failure” stands as a kind of metaphilosophical interlude in a work otherwise devoted to the philosophy of religion. As such, it plays two distinct roles in the larger argument of the lectures. First, it serves to clarify their main thesis. Thus, van Inwagen opens the lecture with the following words:

I have said that my project in these lectures is to defend the conclusion that the argument from evil is a failure. My purpose in the present lecture is to explain what I mean by calling this argument, or any philosophical argument, a failure. (37)

But van Inwagen’s metaphilosophical discussion is also intended to play another role in the larger argument of the lectures. Specifically, and although van Inwagen willingly accepts the burden of showing where exactly the argument from evil fails, he believes that once we are clear about what it would be for a philosophical argument to succeed, we will have good reason to regard it as highly improbable or unlikely that this particular argument is successful. Thus, immediately after presenting his argument for The Pessimistic Verdict (and while temporarily bracketing the argument from evil from its scope), van Inwagen offers the following:

Is it plausible to hold that philosophy can provide a successful argument for the non-existence of God, even though philosophy is unable to provide a successful argument for any other substantive thesis? I have to say that this seems implausible to me. It seems antecedently highly improbable that philosophy, in whose house there have been debated scores (at least) of important questions, should be able to provide a decisive answer to exactly one of them (54).

And he concludes the lecture with the following remark:
My hope is that my reflections on the topic of philosophical argument will lead you to the conclusion that it would be a very odd thing if the argument from evil were a success (55).

While acknowledging the relevance of van Inwagen’s metaphilosophical reflections to the special case of arguments concerning the existence of God, we believe that they are of quite general philosophical interest. In any case, that is the spirit in which we will engage with them.

2. On counterexamples and philosophical substance

Are there any successful philosophical arguments for substantive theses? Let’s begin with a simpler and more straightforward question, albeit one that is still not perfectly transparent: Are there any successful philosophical arguments at all?

Here are three quick candidates, philosophical arguments that at least some philosophers have thought successfully established their conclusions. The first belongs to metaphysics; the second to epistemology; the third to ethics.

(i) In his seminal treatise Material Beings (1990), van Inwagen offers the following argument against compositional nihilism, the view that everything that exists is simple.

(1) I exist.
(2) If I exist, I have proper parts.
(3) Therefore, compositional nihilism is false.3

---

3 We should note that even in the work in which this argument is put forward, van Inwagen’s own opinion of it seems relatively low. There, he includes it in a group of arguments of which he says the following: “These arguments are perhaps rather weak, but I do not think that they are entirely worthless” (1990: 115). Nevertheless, it is still true that there are philosophers who would regard the argument as compelling, even if van Inwagen is not among their number. See, e.g., Schaffer (2009: 358).
(ii) Kelly and McGrath (2010: 332) endorse the following argument:

1. We know that not everyone who is alive today will still be alive fifty years from now.
2. Therefore, we have some knowledge of the future.
3. Therefore, inductive skepticism (the view that we know nothing about the future) is false.

(iii) Consider a standard version of act consequentialism, according to which one is under a standing moral obligation to act so that fewer rather than more people die, in the absence of any further morally significant consequences. Against this theory, someone might offer the following argument:

1. Even if one knows with certainty that (i) one can save the lives of two hospitalized patients dying from organ failure by forcibly overpowering and harvesting the organs of an innocent bystander and that (ii) there is no other way of saving the two patients, one is not morally required to sacrifice the innocent bystander.
2. If act consequentialism is true, one is morally required to sacrifice the bystander.
3. Therefore, act consequentialism is false.

Each of these arguments would be regarded as a compelling argument for its conclusion by at least some philosophers; of course, in each case, there are other philosophers who would deny that it is. (We assume that the mere fact that an argument’s status is *controversial* among philosophers is not enough to settle the question in favor of its detractors.) It is not our purpose to defend any of these arguments in particular; rather, we want to call attention to what they have in common. Each argument purports to refute a substantive philosophical thesis (compositional nihilism, inductive skepticism, act consequentialism) in virtue of providing a counterexample to that thesis. More generally, any putative counterexample to a philosophical thesis can be presented as an argument against that thesis. The claim that there are no successful philosophical arguments thus
amounts to the claim that there are no successful counterexamples to philosophical theses. Put the other way around, anyone who thinks that philosophers at least occasionally devise successful counterexamples should think that philosophers at least occasionally produce successful arguments. (Although of course, the status of specific cases might very well be controversial.) Since we believe that there are some successful counterexamples in philosophy, we believe that there are some successful philosophical arguments.

Of course, The Pessimistic Verdict is not that there are no successful philosophical arguments; rather, it is that there are no successful philosophical arguments for substantive philosophical theses. And indeed, the inclusion of this important qualification seems intended to allow for the successful refutation of philosophical theses, either via the provision of a counterexample or via some other means. Here is van Inwagen’s most extended discussion of the qualification:

I say “substantive philosophical thesis” because I concede that there are, so to call them, minor philosophical theses—such as the thesis that, whatever knowledge may be, it is not simply justified true belief—for which there are arguments that should convert any rational person. I call this thesis minor not because I think that the problem of the analysis of knowledge is unimportant, but precisely because the thesis does not constitute an analysis of knowledge; its message is only that a certain proposed analysis is a failure. Or suppose, as many have supposed, that Gödel’s incompleteness results show, establish that the formalists were wrong about the nature of mathematics. The thesis that formalism is false may in one way be an important philosophical thesis, but only because a lot of people had thought that formalism was true. It is not a substantive philosophical thesis in the way formalism itself is (39).

In a related context elsewhere, van Inwagen writes of “negative” philosophical theses, with reference to the same examples: “If there is any philosophical thesis that all or most philosophers affirm, it is a negative thesis: that formalism is not the right philosophy of
mathematics, for example, or that knowledge is not (simply) justified, true belief” (2004: 334-335).

The idea then, seems to be this: even though Gettier provided successful counterexamples to the justified true belief account of knowledge, and any such counterexample can of course be presented in the form of a successful philosophical argument, the conclusion of that argument is that knowledge is not (simply) justified true belief. And that conclusion is a negative (or “minor”) thesis as opposed to a substantive philosophical thesis. In this respect, it contrasts with the thesis that knowledge is simply justified true belief, which is a philosophically substantive thesis (albeit one that is generally taken to be false in the post-Gettier era). Thus, the fact that there are successful counterexamples in philosophy does not pose a threat to The Pessimistic Verdict, so long as the arguments corresponding to the counterexamples establish only negative or minor philosophical theses.

However, we think that the way in which the ideology of philosophical substantiality is being employed here leads to difficulties. Consider, for example, van Inwagen’s earlier argument against compositional nihilism given above. We assume that compositional nihilism is an example of substantive philosophical thesis in ontology; in this respect, it is like formalism in the philosophy of mathematics or the thesis that knowledge is justified true belief in epistemology. There is some initial temptation then, to say that an argument like van Inwagen’s, which purports to refute compositional nihilism by means of a counterexample, is not itself an argument for a substantive philosophical thesis: even if it is successful, it would establish only the negative or minor philosophical thesis that compositional nihilism is false, as opposed to some philosophically substantial
conclusion. When construed in this way, such arguments pose no threat to The Pessimistic Verdict, for they fall outside of its scope.

The problem, however, is this. There is an obvious sense in which the thesis of compositional nihilism is itself a “negative” thesis: it is the denial of the claim that the universe contains objects with proper parts. An argument that successfully refuted compositional nihilism, either by means of a counterexample or in some other way, would ipso facto be an argument that successfully established a positive thesis to the effect that the universe contains objects of a certain ontological type, viz. objects that are composed of (proper) parts. And why wouldn’t that ontological thesis count as a substantive philosophical thesis? Indeed, that the universe contains objects of such-and-such an ontological type would seem to be a paradigm of a substantive metaphysical thesis. Given this, if van Inwagen’s argument against compositional nihilism is a good one, it is not merely an example of a successful philosophical argument; it is an example of a successful philosophical argument for a substantive philosophical conclusion—and therefore, a counterexample to The Pessimistic Verdict.

A similar situation obtains with respect to the thesis of inductive skepticism in epistemology. On the one hand, inductive skepticism is clearly a substantive philosophical thesis. On the other hand, inductive skepticism is also a negative philosophical thesis: it is the denial of the claim that we possess a certain kind of knowledge. Any successful argument against this negative thesis, via counterexample or some other means, would ipso facto be a successful existence proof that we do have a certain kind of knowledge, knowledge of the future. But the claim that we possess a certain kind of knowledge is itself a substantive claim of epistemology. Thus, any such
argument would itself be a successful argument for a substantive philosophical conclusion, and thus, a counterexample to The Pessimistic Verdict.\(^4\)

One lesson that has emerged so far is the following. Although in some cases, the negation of a substantive philosophical thesis is not a substantive philosophical thesis, in other cases, the negation of a substantive philosophical thesis is itself a substantive philosophical thesis. (Perhaps act consequentialism and the thesis that knowledge is justified true belief are examples of substantive philosophical theses of the former kind, while compositional nihilism and inductive skepticism are examples of the latter.) Thus, one way in which a philosopher could offer a successful argument for a substantive conclusion (and thereby refute The Pessimistic Verdict) is by offering a counterexample to a philosophical thesis that is itself the negation of a substantive philosophical thesis. It is widely thought that, whatever their deficiencies as a group, analytic philosophers are quite good at coming up with counterexamples to each other’s claims. (Indeed, this point is often conceded even by detractors of analytic philosophy, in the course of lamenting its practitioners’ lack of allegedly more valuable intellectual skills.) Is every case in which this has occurred one in which the negation of the refuted thesis is “minor” in van Inwagen’s sense? Perhaps, but we see no reason to think this.

Indeed, when one recalls van Inwagen’s larger dialectical purposes, it seems that there will be significant pressure on him to be extremely inclusive about what counts as a substantive philosophical thesis. Recall from above that one of van Inwagen’s purposes

\(^4\)Is it a relevant difference that while some competent philosophers would deny that the arguments against compositional nihilism and inductive skepticism mentioned above succeed in refuting those views via counterexample, no competent philosophers deny that Gettier succeeded in refuting the justified true belief account of knowledge via counterexample? However, it is not true that all competent philosophers agree that Gettier succeeded in refuting the justified true belief account via counterexample: dissenters include Hetherington (2001) and Weatherson (2003).
in arguing that philosophical failure is ubiquitous is to increase the plausibility of the
claim that the “argument from evil” is itself a failure. Consider, for example, the passage
that immediately follows van Inwagen’s argument for The Pessimistic Verdict:

Now if it is indeed true that no philosophical argument for any substantive conclusion
is successful in the sense that I have proposed, it immediately follows that the
argument from evil is not a success in that sense—given, at any rate, two premises that
I don’t think anyone would deny: that the argument from evil is a philosophical
argument and that the non-existence of God is a substantive philosophical thesis (53).

Here, van Inwagen assumes that everyone would agree that the non-existence of God is a
substantive philosophical thesis. In fact, however, we are inclined to deny this claim,
given his earlier remarks about what it is for something to count as a substantive
philosophical thesis in his sense. After all, the mere denial that God exists does not seem
to have much in common with formalism (understood as a positive account of the nature
of mathematics) or the justified true belief account of knowledge. One can, after all, deny
that God exists without having much in the way of a positive account about the origins or
nature of the universe. In fact, one who argues for the nonexistence of God on the basis
of evil and then declines to say anything more seems in some respects rather like
someone who refutes a proposed analysis of knowledge and then stops before offering an
alternative account in its place. (“Look, I might not know what the true story is here, but
I know that it’s not that...”)

Of course, there is an obvious sense in which the nonexistence of God is a substantive
thesis. Notably, the opposite view has been held (in some times and places virtually
without exception) by many millions of people, including many individuals of the highest
intellectual caliber. So the thesis that God does not exist is hardly a trivial one. But as we
have seen, van Inwagen explicitly denies that the popularity of a philosophical thesis (or
even its popularity among a select group of people) suffices to make the negation of that thesis a philosophically substantive claim. (As we read him, even if all of the most sophisticated epistemologists had been fervent advocates of the view that knowledge is justified true belief prior to Gettier, Gettier’s thesis that knowledge is not justified true belief would still be a “minor” or “negative” thesis as opposed to a substantive one.)

The dilemma that we see for van Inwagen then, is this. Unless he employs some relatively inclusive sense of “substantive”, the negative thesis that God does not exist will fail to count as substantive; rather, like the thesis that knowledge is not justified true belief or that formalism is false, it will be classified as a “minor” philosophical thesis (notwithstanding the considerable violence that this does to the ear). In that case, even if The Pessimistic Verdict turns out to be true, this would have no bearing on the plausibility that the argument from evil is a success; for the truth of The Pessimistic Verdict is consistent with philosophers having produced arguments that successfully establish any number of “minor” philosophical theses (the reference class to which the thesis that God does not exist belongs). On the other hand, if van Inwagen employs a relatively inclusive or liberal sense of “substantive” in order to allow the thesis that God does not exist to qualify (despite its seemingly “negative” content), the harder it is to insist that people like Gettier and Gödel failed to show anything of philosophical substance—an admission which would, of course, immediately refute The Pessimistic Verdict.

3. Van Inwagen’s criterion of philosophical success

The official statement of van Inwagen’s “criterion of philosophical success” runs as follows:
An argument for \( p \) is a success just in case it can be used, under ideal circumstances, to convert an audience of ideal agnostics (agnostics with respect to \( p \)) to belief in \( p \)—in the presence of an ideal opponent of \( p \). (47)

The criterion invokes idealized notions in three distinct places: in its appeal to \textit{ideal circumstances}, \textit{ideal agnostics}, and an \textit{ideal opponent of} \( p \). Following the example of van Inwagen’s own discussion, we will focus primarily on the role played by ideal agnostics.

As characterized by van Inwagen, an \textit{ideal agnostic with respect to} \( p \) has no initial opinion about whether \( p \) is true, and moreover, “no predilection, emotional or otherwise” to accept either \( p \) or not-\( p \) (44). However, the ideal agnostic “would very much like to come to some sort of reasoned opinion—in fact, to achieve knowledge on the matter if possible” (44). Thus, ideal agnostics are \textit{disinterested but not uninterested} with respect to the target proposition. In addition to this orientation, ideal agnostics are “ideally rational” individuals (53), who possess “unlimited leisure and superhuman patience” as well as “the highest possible intelligence and the highest degree of logical and philosophical acumen” (44). On the other hand, an \textit{ideal opponent of} \( p \), while sharing the same general intellectual virtues having to do with rationality, intelligence, and so on, differs in his or her orientation towards the target proposition \( p \): the ideal opponent “will employ every rational means possible, at every stage of the debate, to block the… attempt at conversion” (45).

On one natural reading, van Inwagen’s criterion demands a certain kind of \textit{unanimity}: a successful philosophical argument would be potent enough to convert \textit{all} of the

\footnote{Although van Inwagen says relatively little about what is packed into “ideal circumstances”, the general kind of thing he has in mind seems clear enough from his scattered remarks about the subject. For example, we should not imagine that the attempt at conversion might be short-circuited by running out of time or chalk, or that it is taking place in an uncomfortable or loud environment (cf. p.43). In any case, nothing in what follows will depend crucially on how exactly this notion is understood.}
members of an audience composed of ideal agnostics. One might worry that this sets the bar too high. For example, imagine an argument that converts virtually all of the ideal agnostics to its conclusion (perhaps there is a single, eccentric holdout): wouldn’t it be too harsh to declare such an argument a failure? However, van Inwagen explicitly assumes that the ideal agnostics will respond to any argument in the same way (51-52). Given this assumption, an argument that is strong enough to convert any ideal agnostic will be strong enough to convert them all.6

While further questions of detail could be raised about van Inwagen’s criterion, we wish to develop a quite general objection to the project of characterizing successful arguments in terms of their ability to convert idealized agnostics to their conclusions. Our contention will be that this general framework breaks down when applied to certain core cases, cases that any account of argumentative success should be able to cleanly handle.

Consider first what would seem to be an undeniable case in which a philosopher offers a successful argument. Imagine a philosopher who takes as premises a number of distinct propositions, each one of which is an item of common knowledge, or at least, something that is generally known to be true. From this basis of generally known propositions, she proceeds to prove, by a series of transparently valid steps, some philosophically substantive conclusion— in whatever sense of “philosophically substantive” one prefers. (We can also imagine that each of the premises plays an essential role in the argument: if any one of them were removed, the argument would no

6 Compare the issue that arises for “ideal observer” theories in ethics, about whether anyone who satisfied the requirements for being an ideal observer would converge on the same moral views (Kagan 1998: 274-275). For a recent exchange on the question of whether there is a single, ideally rational response to any body of evidence and arguments, see Kelly (2013) and White (2013).
longer be valid.) We aver that if a philosopher actually succeeded in doing this—that is, in producing a transparently valid argument for a philosophically substantive conclusion from generally known premises—then that argument would be an absolute paradigm of a successful philosophical argument. Thus, it is a condition of adequacy on any proposed criterion of philosophical success that it return the verdict that an argument with these characteristics would be a successful philosophical argument.

However, consider the problem that this kind of case poses for an account which seeks to analyze philosophical success in terms of the ability to convert an audience composed of idealized agnostics. Because the conclusion of the hypothetical argument follows validly from propositions that are generally known, the conclusion follows validly from propositions that are known to members of the audience. Recall next that an essential part of what it is to be an ideal agnostic is to be ideally rational. Now, on standard accounts of ideal rationality, ideally rational agents possess a kind of logical omniscience, and as we have seen, van Inwagen explicitly builds into his characterization of ideal agnostics that they possess “the highest degree of logical acumen”. Given this, it seems that ideal agnostics would not fail to believe propositions that are logically entailed by things that they know. Thus, given that the logically valid argument proceeds from

---

7 To be clear, we do not claim that any actual philosophical argument instantiates this general structure, only that any argument that did would be a success. Indeed, we think that an argument with these characteristics would be something like the holy grail when it comes to philosophical success, at least on one historically prominent model: philosophers have clearly aspired to give valid arguments for philosophically substantive conclusions, and moreover, in the course of doing so they have often deliberately eschewed arguments employing esoteric knowledge as premises in favor of arguments employing premises ostensibly knowable to anyone. (Consider in this context the traditional aspiration to provide valid “proofs” of the existence of God employing only “self evident” premises, as well as Bertrand Russell’s oft-quoted remark that “The point of philosophy is to begin with something so obvious as not to seem worth stating, and to end with something so paradoxical that no one will believe it”). We ourselves do not think that it is a necessary condition on a successful philosophical argument that that argument employ only premises that are common knowledge or generally known, only that it would certainly be acceptable to start with such premises.
generally known premises, *an ideal agnostic would already believe the conclusion of the argument before the argument was ever presented to her*. But this of course contradicts the assumption that the ideal agnostics are actually agnostics, or initially neutral with regard to the truth of the target proposition \( p \). In short, the problem is this: when one attempts to apply the criterion to the kind of paradigm case that we have described, the apparatus of “ideal agnostics” seems to break down, because there is no coherent way of making the members of the target audience both agnostic and ideal. Any member of the audience who was initially agnostic with respect to \( p \) would not be ideal, for she would have failed to believe that which is logically entailed by what she knows to be true. Thus, at least on first inspection, the attempt to explicate the notion of argumentative success in terms of the ability to convert idealized agnostics seems to break down when applied to a core case.

One might try to get around the objection by refining the proposal in various ways. For example, even though the propositions employed as premises in the hypothetical argument are stipulated to be generally known, one might take the relevant test to be how idealized individuals would react to the argument in the absence of any prior knowledge of its premises. Or perhaps, one requires that the ideal agnostics *bracket* their knowledge of the premises, and thus provisionally suspend judgment about whether \( p \). *(Compare Descartes’ professed procedure in the Meditations.)* One then asks: would the subsequent presentation of the argument be enough to convince such individuals that \( p \) is true? If it would, then the argument counts as a success; if it would not, then the argument counts as a failure. However, modified in that way the criterion is clearly unsound: once knowledge of the propositions employed as premises is removed or bracketed, the
presentation of the argument might very well fail to persuade otherwise ideal individuals of the conclusion, but that hardly shows (or even provides reasons to think) that the argument is a failure.

Alternatively, one might try to preserve both the agnosticism and the cognitive ideality of the audience to whom the argument is presented by modifying what is packed into their cognitive ideality. Perhaps we should not assume that the ideal agnostics are logically omniscient, or that they believe all of the logical consequences of what they know. For example, Harman (1986) denies that a fully rational person would believe all of the logical consequences of what she knows, on the grounds that a fully rational person would not clutter her mind by believing things that are logically entailed by what she knows but which are of no interest or importance to her. Notice, however, that qualifying ideality along these lines is unhelpful to van Inwagen’s account. For as noted above, van Inwagen builds into his characterization of the ideal agnostics that they would “very much like to come to some sort of reasoned opinion—in fact, to achieve knowledge if possible” with respect to the target proposition. Thus, a case in which a member of the audience does not believe the conclusion of the argument prior to being presented with it would be a case in which she has failed to believe a proposition that answers a question about which she is concerned to know the truth, despite the fact that the proposition is logically entailed by things that she knows. For this reason, the kind of pragmatic restrictions on full rationality suggested by Harman are not helpful to van Inwagen’s account.

More promisingly, one might insist that the kind of full rationality possessed by the ideal agnostics is consistent with their failing to believe propositions that are logically
entailed by what they know, even when the propositions concern matters of interest. For example, perhaps full rationality does not require one to believe logical consequences of what one knows that are extremely difficult to recognize (cf. Cherniak 1986, McGrath 2010: 66). A given derivation might require one to bring together a relatively large number of seemingly disparate and unrelated pieces of prior knowledge; on this view of rationality, the mere fact that one has failed to believe the conclusion of such an argument before being presented with the argument is consistent with one’s being fully rational. A philosophical argument might thus qualify as successful by converting the ideal agnostics in virtue of making manifest previously unrecognized connections. However, this maneuver seems unpromising as well. For imagine that someone devises a philosophical argument that provides a relatively short or simple proof of a substantive conclusion, one that employs a relatively small number of generally known premises. Surely that argument would qualify as a success. (Indeed, in many ways it would be a more impressive intellectual product—compare the preference among mathematicians for relatively short, elegant proofs where such proofs are available.) If we weaken the kind of rationality possessed by the ideal agnostics enough so that it is consistent with their being initially ignorant of the conclusion even in this kind of case, why suppose that, in some other case in which they are presented with an argument that fails to convert them to its conclusion, this reflects badly on the argument, as opposed to being a manifestation of their own cognitive limitations?

No doubt other maneuvers might be suggested. Indeed, despite what we have said thus far, we do not doubt that, given sufficient fiddling with the notion of cognitive ideality, one could come up with some account that allowed the target audience to be
both agnostic and cognitively ideal in that sense, and which is not subject to the kind of
obvious problems that afflict the proposals that we have considered here. Still, how
much confidence one should invest in the resulting proposal in that event is open to
doubt. After all, the original idea of analyzing philosophically successful arguments as
those that have the power to convert ideally rational audiences of agnostics is an
intuitively attractive one, and an idea that would be illuminating if true. Once one insists
on characterizing the audience in terms of some non-standard account of ideal rationality
or cognitive ideality specially introduced for present purposes, it is doubtful that those
attractions are retained.

Moreover, we think that the problem that we have raised for the proposal reflects
something significant about the practice of attempting to construct compelling
philosophical arguments, and why success in the endeavor is likely to resist analysis in
terms of the conversion of idealized intellects. In paradigmatic cases in which someone
becomes rationally persuaded of a conclusion in virtue of being presented with a
philosophical argument, that process typically involves calling their attention to
connections between propositions (for example, that the relation of entailment obtains
between premises and conclusion) of which they were previously unaware, as opposed to
supplying them with justification or evidence for believing the argument’s premises. Of
course, we can certainly imagine exceptional cases in which the epistemic efficacy of the
presentation of the argument consists in providing justification or warrant for believing its
premises, justification that the members of the audience previously lacked. For
example, imagine a context in which the prevailing norms dictate that someone will
present an argument only if she takes it be sound (so there is no question of someone’s
playing “devil’s advocate”, etc.). Even if a member of the audience was not previously in a position to justifiably believe some premise of the argument, she might know that the person presenting the argument is in a good position to know about such things. In these circumstances, the fact that the person presenting the argument has endorsed it as sound might provide a kind of testimonial justification or warrant for believing the premises; in this way, the audience member might come to justifiably believe the argument’s conclusion. But this is clearly not the usual case, nor the typical aspiration of philosophical argumentation. Rather, in devising a philosophical argument, one selects one’s premises with an eye towards what one’s audience already believes, or is disposed to believe, or already has good reason to believe, and then attempts to show how these propositions support the target conclusion. (Of course, one might at some point realize that some proposition among one’s original premises stands in need of bolstering and attempt to supply reasons for believing it; but at this point, one has strictly speaking begun treating the relevant proposition not as a premise but rather as a lemma in a more complex version of the argument.) In paradigmatic cases in which the presentation of a philosophical argument rationally persuades the audience, the epistemic efficacy of the presentation consists in making manifest the logical or rational connection between premises and conclusion. But that philosophical arguments sometimes play this role presupposes that those to whom they are addressed are less than ideal intellects, for ideal intellects would in effect regard the information conveyed by the presentation of the argument as old news.

On the other hand, standard philosophical practice also presupposes that we resemble ideal intellects in at least this respect: that our beliefs are generally responsive to reasons,
at least when the reasons are called to attention. Just as the practice of offering philosophical arguments would have little if any point among the ideally rational, so too it would have little if any point among the pathologically irrational, who respond with shrugs when it is pointed out that things that they currently believe entail or support things that they do not currently believe. Thus, the ordinary practice of employing philosophical arguments as tools of rational persuasion seems to presuppose that we occupy some intermediate point on the spectrum of rationality, somewhere well above the pathologically irrational but well below the ideally rational.

4. Philosophical disagreement and The Pessimistic Verdict

Recall The Pessimistic Verdict:

There are no successful arguments for substantive philosophical theses.

Let’s waive the line of objection offered in the previous section to van Inwagen’s criterion of philosophical success in order to see how he employs that criterion in arguing for The Pessimistic Verdict. We take van Inwagen’s argument for The Pessimistic Verdict to be contained in passages such as the following:

All philosophical arguments, or at any rate all philosophical arguments that have attracted the attention of the philosophical community, have been tested under circumstances that approximate sufficiently to the circumstances of an ideal debate, that it is reasonable to conclude that they would fail the “ideal debate” test. If any reasonably well-known philosophical argument for a substantive conclusion had the power to convert an unbiased ideal audience to its conclusion (given that it was presented to the audience under ideal conditions), then, to a high probability, assent to the conclusion of that argument would be more widespread among philosophers than assent to any substantive philosophical thesis actually is (53).
Leaving aside those philosophical theses that almost everyone would accept *without argument*, there are no philosophical theses that are both substantive and uncontroversial. If the argument is a success by the terms of my definition, why has it not got the power to produce considerably greater uniformity of opinion among philosophers in the matter of its conclusion than in fact exists? Or if it has got that power, why has this power not been exercised? These questions, I believe, have no good answers (54).

In these passages, the fact that the conclusion of a philosophical argument is typically denied by a significant number of philosophers is treated as compelling evidence that the argument would not convince the ideal agnostics. We concede that how actual philosophers react to an argument is at least *some* evidence as to how the ideal agnostics would react. Of course, how strong that evidence is depends on how closely actual philosophers (and the actual circumstances in which they consider arguments) approximate idealized agnostics (and their idealized circumstances). On the assumption that actual philosophers approximate their idealized counterparts very closely, then the reactions of the former will be comparatively strong evidence about the reactions of the latter; on the other hand, the greater the divergence between actual philosophers and their idealized counterparts, the weaker the evidence is.

Notice that, given van Inwagen’s criterion of success and the level of controversy that we find among actual philosophers, there will tend to be an inverse relationship between (i) one’s optimism about the intellectual qualities of actual philosophers and their circumstances on the one hand, and (ii) one’s optimism about the number of successful philosophical arguments on the other. For example, if one is very optimistic about actual philosophers and their circumstances, and takes them to approximate their idealized counterparts very closely, then one will see the failure of actual philosophers to reach...
agreement as strong evidence that the ideal agnostics would not respond to an argument by converging on its conclusion; one will thus give relatively high credence to The Pessimistic Verdict. If, on the other hand, one is pessimistic about actual philosophers and their circumstances, and tends to see them as a class as falling well short of their idealized counterparts, then one is free to be more optimistic about the kind of convergence that would obtain among the idealized agnostics.

Against this background, a natural worry is that van Inwagen’s pessimism about actual philosophical arguments is an artifact of being overly optimistic about actual philosophers and their circumstances. Offhand, one might think that even the most rational philosophers fall well short of the kind of ideal rationality possessed by the ideal agnostics (however exactly “ideal rationality” is understood). Moreover, even when maximally charitable assumptions are made about the rationality and other intellectual virtues of actual philosophers, there is still the pressing question of how well actual philosophers approximate the initial neutrality of the idealized agnostics. After all, in many cases in which a philosopher is presented with an argument for a substantive conclusion, he or she will already have an opinion about that conclusion. (Notice that in many cases, her having an opinion about the conclusion prior to being presented with the argument will be perfectly reasonable; indeed, in many cases it will be unreasonable for her not to already have an opinion, given her epistemic circumstances.) And in many cases in which she lacks an opinion, she will nevertheless have some “predilection” on way or the other. But anyone who has such an opinion or predilection in a given case already differs from an ideal agnostic in a potentially crucial respect.
Indeed, we think that it can be shown, using assumptions that van Inwagen himself accepts, that the class of actual philosophers is sufficiently dissimilar to the class of ideal agnostics to undermine the kind of plausibility argument that he offers. Consider first philosophical arguments for substantive theses that have in fact convinced large numbers of philosophers of their conclusions, and which are regarded by these philosophers as rationally compelling. We might think here, for example, of Kripke’s case that ordinary proper names are rigid designators (1980), Putnam’s case for semantic externalism (1975), Burge’s argument for “anti-individualism” (1984), or van Inwagen’s arguments for incompatibilism about free will.8 Or consider the philosophical argument that is van Inwagen’s primary concern in the Gifford Lectures, the argument from evil. (Here we will assume, with van Inwagen, that the thesis that God does not exist is in fact a “substantive” one.) Many philosophers believe that the argument from evil is successful. Here, for example, David Lewis, in the posthumously published essay “Divine Evil”:

In my view, even the most ambitious version [of the argument from evil] succeeds conclusively. There is no evasion, unless the standards of success are set unreasonably high. (231).

But of course, there are many philosophers who disagree with this assessment, including van Inwagen himself.

Recall that van Inwagen explicitly assumes that there will be no disagreement among the idealized agnostics: a given philosophical argument will either persuade all of them or

---

8As van Inwagen (2004: 350) notes: “My arguments for incompatibilism…have in fact demonstrated that they have the power to form philosophical opinion: they have convinced some philosophers who were trying to decide whether to be compatibilists or incompatibilists to become incompatibilists”. In an attached footnote, he adds that “I think it is very probable that ‘they have convinced some philosophers’ is a gross understatement. I think it is very probable that they have convinced a great many philosophers” (emphasis added). We agree.
else none of them. Now, if Lewis is right that the argument from evil “succeeds conclusively” and can be resisted only if “the standards of success are set unreasonably high”, then clearly, the ideal agnostics would be persuaded by the argument. (Presumably, the ideal agnostics would not insist on unreasonably high standards of success before coming to believe the conclusion.) And again, there are many generally reasonable actual philosophers who share Lewis’ assessment. On the other hand, if van Inwagen’s assessment of the argument from evil is correct, then the ideal agnostics would remain unpersuaded. (And of course, there are many generally reasonable philosophers who share van Inwagen’s assessment of the argument.) What we can conclude from this is the following: given that there is some fact of the matter about what the ideal agnostics would think about the argument from evil, there are many generally reasonable philosophers whose own opinion about the argument is a very bad guide to what that fact of the matter is. Moreover, the same is true, mutatis mutandis, of Kripke’s arguments that ordinary proper names are rigid designators, Putnam’s arguments for semantic externalism, van Inwagen’s arguments for incompatibilism, and any number of other arguments known to the philosophical community. In view of this, the level of optimism as to how closely actual philosophers approximate their idealized counterparts required by Inwagen’s argument seems unwarranted: if there is one thing that we can be sure of in this vicinity, it’s that there is a tremendous amount of slack between the assessments of actual philosophers (given the kinds of disagreements we find about whether particular arguments are compelling) and the assessments of their idealized counterparts (given their complete lack of disagreement).
Consider next the class of philosophical arguments of which the following is true: informed philosophical opinion is substantially divided between those who regard the argument as rationally compelling, and those who deny that it is. In these cases, is there any general reason to think that someone who makes the substantive judgment that an argument is *not* compelling is more likely to be correct, or tracking the view of the ideal agnostics, than someone who makes the opposite substantive judgment? As far as we can see, there is not. Still less is there any reason to suppose that those who make the substantive judgment that an argument is not compelling are tracking what the ideal agnostics think *in every case*, which is what The Pessimistic Verdict requires. It is not, after all, as though Lewis’ substantive judgment that the argument from evil is rationally compelling (and that one who responds to the argument by believing its conclusion is responding in the most reasonable way) is more likely to be wrong than van Inwagen’s opposite substantive judgment simply because of the *type* of judgment that it is. To the extent that we have good reasons to believe that the argument from evil is a failure, this is due to the kind of direct engagement with, and meticulous, sustained critique of the argument that van Inwagen provides in the rest of the lectures, as opposed to the kind of general considerations about philosophical failure considered here.⁹

5. Conclusion

It is occasionally remarked that there is something distinctively odd about the genre of philosophical *festschrift*. In order to honor an unusually distinguished living member of the discipline, a group of that person’s scholarly peers, past and current colleagues, 

⁹For more general reflections on the epistemic significance of persistent disagreement among actual philosophers, see Kelly (forthcoming).
former students, and admirers are invited to contribute essays on some aspect of his or
her life’s work. In their essays, the contributors typically set out to show just how
wrongheaded the distinguished philosopher’s views on some topic really are, or at least,
that his or her arguments on that topic are unsound and ought not to convince anyone. In
an obvious way, our own contribution to the current volume is in keeping with this rather
questionable tradition. For we have attempted to show that, with respect to the topic of
philosophical success and failure, van Inwagen’s arguments are unpersuasive.

There is another respect, however, in which our contribution can be viewed as more
in keeping with the true, underlying spirit of the occasion. For decades, Peter van
Inwagen has been at the forefront of philosophy. Few living philosophers can match the
power, depth and influence of his writings across a comparable range of areas. Moreover,
unlike many philosophers of significant influence—and we take this to be greatly to his
credit—van Inwagen’s influence is due in no small measure to the perceived power of his
arguments. Now, as van Inwagen is well aware, if his metaphilosophical views about
success and failure in philosophy are correct, then it is overwhelmingly likely that all of
these arguments are failures. As he writes:

...are there any successful philosophical arguments? I know of none...I hate to admit
this, if only because I should like to think that some of the arguments associated with
my name are successes. But I have to admit that it’s at best highly improbable that
they are (52).

However, if we are correct, and van Inwagen’s metaphilosophical views are mistaken,
then it is a distinct possibility that some of the arguments associated with his name are, in
fact, genuine successes, notwithstanding the fact that some among us remain stubbornly unconvinced.\textsuperscript{10}

\textsuperscript{10} Earlier versions of some of these ideas were presented in a graduate seminar co-taught with John Hawthorne at Princeton University in the spring of 2013; we would like to thank John and the students for their feedback on that occasion. Thanks also to David Chalmers for reading and commenting on an earlier draft, to John Keller for extending an invitation to contribute to this volume, and (last but not least) to Peter van Inwagen for his stimulating writings over the years, from which both of us have learned much.
References


