

A Counterexample to van Inwagen's Criterion for the Success and Failure of Philosophical Arguments

By Daniel M. Johnson
Baylor University, 2007

Peter van Inwagen, in his recently published Gifford lectures, proposes the following criterion for philosophical success: “an argument for p is a success just in the case that 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 to p .”¹ An ideal agnostic “will have no initial opinion” with respect to the issue in question and (this is key) “no predilection, emotional or otherwise” one way or the other.

I want to challenge the idea that, for a philosophical argument to count as a success, it must be able to persuade someone that has no already-existing tendency to accept the conclusion.² I'll do this by describing a counterexample, a type of argument that can be *rationaly persuasive* and yet actually depend for its force upon a natural human tendency to accept the conclusion. Since I don't know what more we can demand from a philosophical argument than that it be rationally persuasive, it follows that a successful philosophical argument need not meet Dr. van Inwagen's criterion.

The Counterexample

The argument's form is as follows. First, the person to whom the argument is being given (the audience member) is immediately (non-inferentially) justified in believing the conclusion – so that her justification cannot be put into argument form. However, this audience member does

¹ Peter van Inwagen, *The Problem of Evil* (Oxford: Oxford University Press, 2006), 47.

² The presence of an ideal opponent who can give opposing arguments is not immediately relevant for my purposes; if the type of argument I describe can rationally persuade someone, I think it can rationally persuade even in the face of opposition.

not, for whatever reason, actually believe the conclusion. Perhaps she was persuaded at some point by a fallacious argument, so that she no longer accepts her immediate justification. Second, it is a strong natural human tendency to believe the conclusion of the argument, so that it is difficult to consistently doubt it. So this audience member may very well have failed to remove from her web of belief all of her beliefs which depend for their justification on the questioned belief. Third, the justification of one of the premises depends on a preexistent justification to believe the conclusion – the audience member is only justified in believing the premise if she is already justified in believing the conclusion.

I'd like to make two points about arguments of this type. First, the argument likely will *not* persuade someone with no natural tendency to believe the conclusion. Such a person will likely not accept the premise, because she will not have any beliefs that depend for their justification on the doubted belief. She is a consistent doubter. So this argument will not be able to persuade van Inwagen's audience of ideal agnostics.

The second point, though, is that this argument likely *can* persuade someone who *does* have a natural tendency to believe the conclusion. Such a person may find it extremely difficult to consistently doubt the contested belief, and may retain many beliefs that depend for their justification on the contested belief. So she may well accept the premise (even though its justification depends on her being justified in a belief that she doesn't have), and upon seeing that the premise entails the conclusion, may well be moved to accept the conclusion.³

Now, one immediate response would be to deny that the persuasion effected by this argument is rational – after all, isn't it irrational to accept the premise, given that its justification actually depends on the conclusion's being justified? I acknowledge that this argument cannot

³ Now, this may be effective even in the face of opposition, so I think this example holds even in the context of a debate.

give the audience member *new* justification for believing the conclusion – she is not justified in believing the conclusion *on the basis of* the argument. However, recall that she is in fact justified in believing the conclusion – *non-inferentially* justified. So there are actually two irrationalities in the situation: she is irrational in accepting the premise in light of her doubt of the conclusion, but she is also irrational in doubting the conclusion in the first place. Her acceptance of this argument actually removes *both* irrationalities. Her acceptance of the conclusion renders her rational in believing both the conclusion and the premise. So, I conclude, this argument can indeed be *rationally* persuasive.

Here is a more concrete example. On some interpretations of what is going on in G.E. Moore's proof of an external world, Moore's proof is an example of the type of argument I am describing. Moore's proof goes like this:

- (I) Here is a hand.
- (II) If hands exist, external things exist.
- (III) Therefore, external things exist.

Some have argued that this proof is circular, since I am only justified in believing (I) if I am already justified in believing (III), and I am non-inferentially justified in believing (III).⁴ So this proof cannot give me new justification to believe (III). I would add, however, that this argument can still be rationally persuasive according to the pattern I presented earlier. I have a strong natural tendency to believe (III), and so even when I doubt it I may not doubt (I) even though my justification for believing (I) depends on my justification for believing (III). So, upon being

⁴ See, for example, Crispin Wright, "(Anti-)Skeptics Simple and Subtle: G.E. Moore and John McDowell," in *Philosophy and Phenomenological Research* 65 (2002): 331-49; and Ram Neta, "Fixing the Transmission: The New Mooreans," in *Themes from G.E. Moore: New Essays in Epistemology and Ethics*, ed. by Susana I. Nuccetelli and Gary Seay (Oxford: Oxford University Press, forthcoming).

presented with Moore's proof, I may be persuaded, and this would remove both my irrationality in doubting (III) and my irrationality in accepting (I) in light of my doubt of (III).⁵

My concern, then, is this: does it really always make sense to require the "ideal observer" to have no predilection whatever to accept the conclusion of an argument? I have given an example, I think, of a rationally persuasive (i.e., successful) philosophical proof that actually depends for its force upon a natural human tendency to accept the conclusion. I even think that many arguments of this type, depending on the strength of the particular natural human tendency in question, will be persuasive even in the face of the opposition of the ideal opponent. Because of this, we cannot always, in evaluating an argument, remove all natural human tendencies to accept the conclusions of those arguments.

An Objection

I'd like to take a brief look at one possible objection. For my counterexample argument to work, it must be the case that the audience member is immediately justified in believing the conclusion of the argument, has a natural tendency to do so, and still (irrationally) does not have the belief. Dr. van Inwagen may point out that the ideal observers that he described are never irrational – they never withhold belief when they are justified in believing. So, he may say, arguments of this type could never persuade an ideal observer, since the ideal observer would already have the belief.

I concede that the kind of argument I have described could not persuade Dr. van Inwagen's ideal observers – that is why it is a counterexample to his criterion. There is a stronger

⁵ Others, like Jim Pryor, have argued that I am immediately justified in believing (I), and that my justification for (I) does not depend on my justification for (III). If he is right, then Moore's proof would not be an instance of the type of argument I am describing. It would be easy enough to describe another example; the point of using Moore's proof is its familiarity to many. See Jim Pryor, "What's Wrong with Moore's Argument?" in *Philosophical Issues* 14 (2004): 349-78.

point to be made here, though. It is true that the ideal observers, if they are perfectly rational, would never withhold belief from some proposition they are non-inferentially justified in believing. This, however, would make any philosophical argument for an immediately justified proposition impossible. There would be no way to evaluate such an argument – since there would be no audience of “ideal agnostics” to direct the argument to (since any ideally rational person would not be agnostic), the argument wouldn’t be either a success or a failure (perhaps it wouldn’t count as an argument at all). This is especially troubling, since it would not only rule out the circular arguments I have described, but all arguments (even normal, non-circular ones) for immediately justified beliefs.

This is another problem for the very conception of an “ideal agnostic.” Why should we demand that a philosophical argument persuade “ideal agnostics” at all? Why can’t arguments that rationally persuade non-ideal, regular folks count as philosophical successes? I have described an argument that can only rationally persuade non-ideal agnostics, but the problem of evaluating arguments for immediately justified beliefs gives, I think, yet another reason to accept that we should not use the concept of the “ideal agnostic” to evaluate arguments.

Concluding Remarks

I’d like to make a few concluding observations about the consequences for Dr. van Inwagen’s views. First, Dr. van Inwagen understands that any criterion for philosophical success needs to be relative to some context – this is why he posits the whole “ideal debate” idea in the first place. He thinks that philosophical arguments should be evaluated in some way by their ability to persuade people, not just by some abstract characteristic of the argument in isolation from those who are arguing. I enthusiastically agree with the need for context-sensitivity in

evaluating arguments, but I would argue that it may be that such a context just cannot be idealized. The fact is that we are human beings, with irrationalities and natural tendencies. I think we should recognize that some arguments actually are able to rationally persuade us *because* of those irrationalities and natural tendencies. Whether an argument is philosophically successful, then, will depend on the actual context in which it is given – the particular situation of the human being to which the argument is directed.

Second, I'd like to bring my objection to Dr. van Inwagen's criterion for philosophical success into more direct contact with his main line of argument, his evaluation of the atheological argument from evil. If Calvin is right about the sense of deity that all people possess, it may well be that the proposition "God exists" is precisely the type of belief that can serve as the conclusion for the type of argument I presented at the beginning: all people are non-inferentially justified in believing that God exists, have a strong natural tendency to believe it, but sinfully (irrationally) suppress the belief. It may be possible, then, to construct rationally persuasive arguments for the conclusion that God exists with premises that depend for their justification on a preexistent justification to believe in God.

In any case, I think it is a mistake to use idealized versions of human beings to evaluate philosophical arguments, because some arguments that depend on non-ideal factors for their ability to persuade should still be counted as philosophical successes.