

A Diagnosis for the Circularity of Anti-Skeptical Arguments

RACHEL BAYEFSKY
Yale University

Arguments against skeptical hypotheses often raise the concern that their premises arise through a method that would be unreliable if the skeptical hypothesis turned out to be true. Such arguments may be considered guilty of circularity if their premises rely on the falsity of the skeptical hypothesis in an inappropriate manner—a manner I intend to clarify. In this paper, I plan to develop a diagnosis for the circularity of anti-skeptical arguments. The diagnosis, which I will call D, will be partially based on Robert Nozick’s “subjunctive conditionals” account of knowledge. I will apply D to various anti-skeptical arguments. In particular, I intend to use D to contend that Descartes’ proof of the truth of his clear and distinct perceptions is not necessarily circular.

I will argue that D provides an explanation for the way in which the effectiveness of an argument is crippled by circularity. The effectiveness of an argument is its ability to take people who are committed to evaluating arguments based on rational standards and bring them closer to the conclusion than they are when they first encounter the argument. An effective argument in this sense is not one that persuades by rhetorical force, but one whose logical steps make progress towards the conclusion. Circularity impedes an argument’s effectiveness because the conclusion must be accepted before the premises, which means that the argument itself is not responsible for advancing the conclusion. Diagnosis D, by clarifying the reasons why the conclusion of a circular argument against a skeptical hypothesis must be accepted before at least one of the premises is accepted, provides us with a way to explain the decrease in the effectiveness of an argument due to circularity.

Here are some examples of arguments that might be considered circular. First, let us say that I am suddenly seized with the worry that the floor of my building, instead of being connected solidly to the ground, is floating on air. I try to reassure myself with the thought that, if I wanted, I could walk out of my room right now and go down the elevator in the ordinary way. Here is my “argument:”

(1-e): If the floor of my building has suddenly begun to float on air, I cannot walk out of my room and go down the elevator in the ordinary way.

(2-e): I can walk out of my room right now and go down the elevator in the ordinary way.

(C-e): Therefore, the floor of my building has not suddenly begun to float on air.

Alternatively, suppose a congenitally blind person receives surgery and slowly begins to recover her ability to experience visual imagery.¹ She sees the outline of a car in front of her, considers the situation, and reaches the fortunate conclusion: “My blindness has been at least partially cured!” Her “argument” runs as follows:

(1-b): If my congenital blindness has not been partially cured, I cannot see the outline of a car in front of me.

(2-b): I am able to see the outline of a car in front of me.

(C-b): Therefore, my blindness has been partially cured.

Are the “elevator” and “blindness” arguments circular? Not in an obvious way: neither conclusion appears explicitly among their respective premises. Clarifying the question of whether and how such arguments are circular is not a simple task. One response to the “elevator” argument, for instance, might be to object that if the conclusion is false, premise (2-e), at least, is also false: if the floor of my building has suddenly begun to float on air, I am unable to walk out of my room and take the elevator down. But for all valid deductive arguments, if the conclusion is false, one or more of the premises must be false. This does not imply that the truth of a premise “relies on the truth of the conclusion” in an inappropriate manner. Take, for instance, the classic argument:

(1-s): All men are mortal.

(2-s): Socrates is a man.

(C-s): Therefore, Socrates is mortal.

Here, the falsity of the conclusion would necessarily undermine the truth of at least one of the two premises. If the “Socrates” argument is circular, however, it is difficult to see how many valid deductive inferences could still stand.

I. A DIAGNOSIS OF CIRCULARITY

My diagnosis for the circularity of an argument will focus on a person's reasons for accepting the premises of the argument and whether or not they would be replicated in the case that the conclusion is false. In formulating my diagnosis, I will adapt one of the conditions of Robert Nozick's account of knowledge in *Philosophical Explanations*. Nozick presents a "subjunctive conditionals" account of knowledge, according to which S knows that p if and only if:

- (1) p is true:
- (2) S believes that p:
- (3) If p were not true, S wouldn't believe that p; and,
- (4) If p were true, S would believe that p.²

The idea behind conditions (3) and (4) is that one's beliefs should only count as knowledge if they are sensitive to the truth. That is, if one would believe p whether or not p is actually true, one's beliefs do not have the epistemic status of knowledge. Here, Nozick is presenting an account of knowledge, not putting forward criteria for evaluating circularity. His third criterion for knowledge, however, can be adapted to form part of my diagnosis D for circular arguments against skeptical hypotheses. According to D, an argument against a skeptical hypothesis is circular with respect to a specific premise p³ if and only if the following conditions hold:

D1: If the conclusion were false, S would still believe that p for the same reasons as if the conclusion were true.

D2: If the conclusion were false, p would be false.

Together, these conditions entail the condition "If p were false, S would still believe that p for the same reasons as if the conclusion were true." While we might simply use this condition as an indication of circularity, it is useful to break it down into D1 and D2 in order to see why different types of arguments beg the question.

The basic idea behind D is that a circular argument, in the most general form, is one whose conclusion is contained in the premises of an argument. But it is not necessary for a conclusion to be explicitly present among the premises in order for an argument to be circular. Rather, an argument against a skeptical hypothesis is circular if the truth of the conclusion must be accepted before the reasons for at least one of the premises can count as justification for the truth of that premise. If I have the same reasons for believing the premise regardless of whether the conclusion is true, and the truth of the premise depends on the truth of the conclusion, then my reasons for

believing the premise do not point me towards the truth of the premise unless I am already assured of the truth of the conclusion. So the purpose of D is to evaluate the strength of the reasoning behind the acceptance of the premise. D explains that an argument against a skeptical hypothesis is ineffective because the conclusion must already have been accepted in order for the premises to be evaluated. Such an argument does not succeed in using the premises to prove the conclusion.

Before testing D against examples of potentially circular arguments, I would like to explain why D1 includes the clause “for the same reasons,” as in, “if the conclusion were false, S would believe that p *for the same reasons* as if the conclusion were true.” The addition of “for the same reasons” rules out a situation in which S would believe that p for different reasons than if the conclusion were false. In such a scenario, S’s reasons could still serve as independent evidence for p even though S would believe that p if the conclusion were false, because S might be able to distinguish the types of reasons available if the conclusion is true from the types of reasons available if the conclusion is false. For instance, let us say the conclusion of an argument is that aliens have landed on earth. One of the premises of the argument is that three aliens have been spotted in a field in Northern Ontario. If the conclusion is true, S’s reasons for believing the premise might include a news report; if the conclusion is false, S’s reasons for believing the premise might instead be limited to a dramatic dream-vision. If the conclusion were false, therefore, S would believe that three aliens had been spotted for a different reason than if the conclusion were true. S might well be able to use the differences in his reasons for believing the premise to evaluate the truth of the conclusion. The fact that S would fail to meet the criterion laid out in D1 reflects the fact that his argument is more effective than an argument judged to be circular under D. This is because S would not necessarily have to accept the conclusion in order to evaluate the evidence for the premise.

II. EVALUATING THE CONDITIONAL

How are we to understand the conditional that is D1? One way of answering this question is to appeal to the possible-worlds account of subjunctive conditionals. We can follow the example of Nozick, who explains the application of his third criterion for knowledge as follows:

“What the subjunctive 3 [if p were false, S wouldn’t believe that p] speaks of is the situation that would hold if p were false. Not every possible situation in which p is false is the situation that would hold if p were false. To fall into possible worlds talk, the subjunctive 3 speaks of the not-p world that is closest to the actual world, or of those not-p worlds that are closest to the actual world, or more strongly (according to my suggestion) of the not-p neighborhood of the actual world. And it is of this or these not-p worlds that it says (in them) S does not believe

that p. What happens in yet other more distant not-p worlds is no concern of the subjunctive 3.”⁴

In other words, in order to evaluate whether the subjunctive conditional 3 holds, we identify the possible world or worlds that are closest to the actual world but that differ from the actual world in the sense that p does not hold. Then we determine whether S believes that p in the closest not-p world or worlds. Applied to D1, in order to evaluate whether S would believe that p if the conclusion were false, we must identify the possible world or worlds closest to the actual world in which the conclusion is false and then determine whether S believes that p in these worlds.

The concept of a “closest possible world or worlds” is, of course, problematic. How do we determine which possible world is closest? Can there be two possible worlds which are different from each other, but equally “close” to this world? The coherence of the “closest possible world” concept has been discussed extensively in the literature on Nozick and others, and I cannot provide a complete solution of this problem here. In general, however, the “closest possible world” is one in which we see the fewest changes from the actual world in terms of physical laws, objects and people that are present, and historical trajectory. Since it is possible for some of these conditions to change while others remain the same, it may be hard to determine, especially in close cases, which worlds count as “closer” than others. The difficulty of choosing the closest possible world when two worlds are rather similar, however, should not preclude us from acknowledging that the distinction is meaningful when applied to possible worlds that are radically different. An example may be instructive. Let us say I am sitting in a park on an ordinary day. Now let us think of possible worlds in which it is false that I am sitting in the park. In possible world W1, I am not sitting in the park because I decided to go to the grocery store instead of the park that morning. In possible world W2, I am not sitting in the park because I have been pulled up into the sky by a reverse gravitational force, which took effect only for one instant and only in one place on earth—the very park bench where I was sitting.

It seems plausible to suggest that W1 is closer to the actual world than W2. This is because although W1 differs from the actual world in that I am no longer sitting in the park, it differs in a way that is predictable and routine from the perspective of someone sitting in the actual world. W1 involves, for instance, no reversals of the laws of nature of the kind that take place in W2.

The notion of possible worlds on which I will be relying is comparative. I will be determining which of different possible worlds is closest to the actual world, not attempting to delineate the exact features of the possible world that is absolutely closest to the actual world. Even without an exhaustive list of characteristics that would make a possible world closest to the actual world, it is possible to determine which of a selection of worlds is most likely to be closest. But the notion of the closest possible world, and the diagnosis in general, must be clarified through examples.

III. TESTING THE DIAGNOSIS

Let us return to the “elevator” example, the argument against the skeptical hypothesis that my floor of the building is floating on air:

(1-e): If my floor of the building has suddenly begun to float on air, I cannot walk out of my room and go down the elevator in the ordinary way.

(2-e): I can walk out of my room right now and go down the elevator in the ordinary way.

(C-e): Therefore, my floor of the building has not suddenly begun to float on air.

We must evaluate both premises of the argument to determine whether or not D1 and D2 hold. Let us consider what would happen if (C-e) were false. If my floor of the building had suddenly begun to float on air, I would be likely to fall to the ground as soon as I walked out of my room. It seems almost impossible for my floor of the building both to be floating on air and to be connected solidly to the ground by an elevator shaft. In the closest possible world in which (C-e) were false, then, (1-e) would be true. I could not simply ride down the elevator. So D2 does not hold with respect to (1-e). Now we diagnose the argument’s circularity with respect to premise (2-e). If my floor of the building has suddenly begun to float on air, I have the same reasons for believing that I can walk out of the room and go down the elevator as I would have had if my floor were as solidly connected to the rest of the building as ever. I have gone down the elevator many times before, I have no particular reason to suppose a great anomaly in the earth’s physical laws has occurred, and so on. So D1 holds with respect to (2-e). Moreover, if (C-e) is false, (2-e) would also be false; if my floor of the building is floating on air, I cannot go down the elevator in the ordinary way. Both D1 and D2, then, hold for the elevator argument with respect to premise (2-e), and so the elevator argument is circular according to diagnosis D.

How does the diagnosis of circularity point out the “elevator” argument’s ineffectiveness? All of my evidence for believing the premise “I can walk out of my room and go down the elevator” is compatible with the premise’s falsehood. This premise will only be plausible if I already believe that my floor is not floating on air. But if I already believe that my floor is not floating on air, the anti-skeptical argument has not been responsible for bringing me any closer to the conclusion that my floor is not floating on air. If I am only soothed by rational arguments, then the “elevator” argument will not be particularly soothing.

Secondly, let us test D on the “blindness” argument:

(1-b): If my congenital blindness has not been partially cured, I cannot

see the outline of a car in front of me.

(2-b): I am able to see the outline of a car in front of me.

(C-b): Therefore, my blindness has been partially cured.

The truth of (1-b) follows from the nature of congenital blindness at this stage of scientific research. Congenitally blind people are widely thought to be unable to experience visual imagery at any time, including in their dreams.⁵ So D2 would not hold of (1-b), because (1-b) is never false. The question is therefore whether or not the “blindness” argument is circular with respect to premise (2-b). There are multiple scenarios under which a woman’s blindness has not been partially cured. One such scenario is a situation in which this woman seems to be able to see but has not actually been cured of blindness. Say the woman is suddenly abducted by aliens and made into a bodiless brain in a vat (BIV) with no eyes at all. It would be possible for the aliens running the BIV system to stimulate the neurons responsible for sight in the woman’s brain. The aliens could give her an experience of seeing the outlines of a car without her blindness actually having been cured. As soon as the aliens withdraw their tentacles, affix the woman’s brain to her body once more, and let her go, the woman will be as blind as ever.

If we turn to the BIV world when we consider what happens when the conclusion of the woman’s argument (“My blindness has been partially cured”), then the woman’s argument would be circular. D2 would apply if the woman is not actually seeing the outlines of the car. This would be because “seeing” requires some causal connection between an external source of visual images and the eye. Though the BIV experience provides images of the outline of a car to the blind woman, these images correspond to nothing in the outside world, and so the woman is not actually seeing the outlines of a car. D1 would hold because the woman would have the same reasons to believe she is seeing the outlines of a car whether or not her blindness is partially cured: her sensations and ideas of the car would be the same. So the woman’s argument that her blindness has been at least partially cured is circular under D if the closest possible world is the BIV world.

But why should we base our judgment of whether D1 holds in the possible world in which the woman is a BIV? As Nozick points out, the subjunctive conditional “if p were false then S wouldn’t believe that p” can be true “even though there is a possible situation where not-p and S believes that p,” are both true, because “not every possible situation in which p is false is the situation that would hold if p were false.”⁶ Regardless of whether D1 would hold in all of the situations in which (C-b) is false—and it would hold in the BIV world—a *closer* possible world in which (C-b) is false is a world where the surgery was simply unsuccessful. In this world, if (C-b) were not true, the woman would not believe (2-b). If the woman’s blindness were not partially cured, she would not think she could see the outlines of a car. D1, therefore, does not hold, and the “blindness” argument is not circular according to D. This is a consequence of the fact

that the scenario in which the woman is a BIV being fed images by aliens is farther away from the actual world than the scenario in which the surgery simply did not work.

The fact that the “blindness” argument is not circular under D provides us with a reason why it is more effective than the “elevator” argument. The congenitally blind woman receives sensory impulses that she recognizes, based perhaps on the testimony of others, as sight. She plausibly supposes that the most likely scenario under which these impulses appear to her is the scenario in which she can actually see. She does not have to believe that her blindness has been cured in order to realize that she is actually seeing; rather, she uses the fact that she is actually seeing to deduce that her blindness has been cured. The argument is therefore more effective than the “elevator” argument because it has brought the woman closer to its conclusion than she was when she began considering the premises.

IV. APPLICATION TO THE CARTESIAN CIRCLE

Now that I have provided a diagnosis of circular arguments and applied it to the “elevator” and “blindness” arguments, I would like to apply my diagnosis to the Cartesian Circle. I intend to show that the Cartesian Circle is not necessarily circular under my diagnosis. In the *Meditations*, Descartes seeks to argue for the conclusion that as a rule, his clear and distinct perceptions are true.⁷ One potential analysis of Descartes’ argument is Keith DeRose’s interpretation, according to which Descartes argues as follows:⁸

(1): For certain propositions, such as the cogito, $2+3=5$, and the causal principle of ideas (i.e. that there must be “at least as much reality in the cause as in its effect”)⁹, if I clearly and distinctly perceive these propositions, then they are true.

(2): **God exists and is no deceiver.

(3): I clearly and distinctly perceive the Rule of Truth: for all p, if I clearly and distinctly perceive that p, then p is true.

(C): All of my clear and distinct propositions have the status of *scientia*, that is, certain knowledge that cannot be doubted.

The ** represents Descartes’ argument for the existence of a non-deceiving God. According to this presentation of Descartes’ argument, before Descartes recognizes the existence of a non-deceiving God, he can attain a fairly high level of certainty regarding his clear and distinct perceptions. But even a small doubt of the general rule of truth can undermine Descartes’ acceptance of (1) by casting doubt on the reliability

of the belief-forming mechanism that results in (1): clear and distinct perception. Descartes must therefore prove the existence of a non-deceiving God in order to clearly and distinctly perceive the general rule of truth, which tells him that as a rule his clear and distinct perceptions are true. When Descartes reaches clear and distinct perception of the general rule of truth, all of his clear and distinct perceptions are raised to a higher level of certainty called *scientia*. The advantage of *scientia*, as opposed to mere clear and distinct perception, is that it is impervious to skeptical attack arising due to doubt of the general rule of truth.

We can now examine whether this two-level version of Descartes' argument exhibits circularity according to D. I will concede Descartes' argument for the existence of a non-deceiving God based on the clear and distinct perceptions in (1).¹⁰ I will also accept the link between (3) and "C", because the reasons for why (3) leads to "C" are contained in the idea of *scientia*. In order to compare this argument to the others I have considered, I will reformulate it as follows:

(1-d): If I clearly and distinctly perceive certain propositions, then they are true. These propositions include $2+3=5$, the cogito, and the causal principle of ideas.

(2-d): There exists a non-deceiving God who can guarantee the truth of my clear and distinct perceptions.

(C-d): I clearly and distinctly perceive the Rule of Truth: for all p, if I clearly and distinctly perceive that p, then p is true.

Since I have conceded Descartes' argument for a non-deceiving God, D2 does not apply with respect to (2-d). Whether or not (C-d) is false, I will have to concede that (2-d) is true. Since the argument is not circular under D with respect to (2-d), I will ask whether the argument is circular under D with respect to (1-d). To test D1, we ask: if (C-d) is false, would I still believe (1-d) for the same reasons as if (C-d) were true? To test D2, we ask: if (C-d) is false, would (1-d) be false?

The question about D1 can be answered in the affirmative. Even if it is not the case that as a rule, every single one of my clear and distinct perceptions are true, I would still believe that I clearly and distinctly perceive certain propositions to be true. The closest possible world in which the general rule of truth is false is not the world ruled by an utterly vindictive evil genius who ensures that all of my clear and distinct perceptions are false. A closer possible world in which (C-d) is false is one in which some, but not all, of my clear and distinct perceptions are true, including those I seem to perceive most clearly and distinctly, such as "I exist." In such a world, I would indeed have the same reasons for thinking "I exist" is true; it would seem impossible for me to doubt this proposition, it would conform with everything I have experienced so far, and so on. D1, in other words, would be satisfied: if (C-d) were false, I would still have the same reasons for believing (1-d) as if (C-d) were true.

D2, however, would not be satisfied. Even if the general rule of truth were false, it does not have to be the case that all of my clear and distinct perceptions are false. In the closest possible world in which the general rule of truth, (C-d), is false, the proposition “I exist,” and likely some other clear and distinct perceptions, could still be true. It is then not the case that the falsity of (C-d) implies the falsity of (1-d). My reasons for believing in the truth of certain clear and distinct perceptions such as “I exist,” though these reasons would be present even if the rule of truth were false, are still able to support (1-d) independently of the falsity of the general Rule of Truth. Descartes’ argument on DeRose’s interpretation, therefore, is not circular according to D.

If we concede Descartes’ argument for the existence of a non-deceiving God, diagnosis D can explain the effectiveness of Descartes’ argument for the truth of his clear and distinct perceptions. Descartes can accept the truth of some of his clear and distinct perceptions without having previously accepted that all of them, as a rule, are true. So he can use the initial clear and distinct perceptions as springboards to prove the existence of God and thereby the truth of all his clear and distinct perceptions. The argumentative steps in Descartes’ proof can bring someone closer to the general Rule of Truth than he or she was at the beginning of the argument. Descartes does not have to assume a world of widespread clear and distinct knowledge in order to work his way up to such a world. His argument, diagnosis D shows us, is thereby rendered more effective.

In this paper, I have sought to provide a diagnosis for the circularity of arguments against skeptical hypotheses and to argue that Descartes’ argument for the truth of his clear and distinct perceptions is not necessarily circular. I suggested that the “elevator” argument was circular under D because my reasons for supposing I could walk outside and take the elevator only counted as reasons for this premise provided the conclusion of the argument—that my floor was not floating on air—was true. I also argued that the “blindness” argument was not circular under D because the woman can believe she is able to see for reasons that would not be available in the closest possible world in which she is blind: a world in which her continuing blindness is due to failed surgery instead of her being a BIV. In the “blindness” but not the “elevator” argument, the premises can be accepted on the basis of the evidence for them without previous commitment to the conclusions. The premises of Descartes’ argument for the truth of his clear and distinct perceptions, in its two-level variety, can also be accepted independently of the conclusion, for even if the conclusion does not hold, the premises—individual clear and distinct perceptions—could still stand.

I have also suggested that D provides a way to explain why circularity detracts from an argument’s effectiveness. In a circular argument, one cannot accept at least one of the premises without also accepting the conclusion, but according to the logic of the argument, one should not be accepting the conclusion without accepting the premises. One therefore has no reason to accept either the premises or the conclusion unless one has already accepted both—but if one has already accepted both, the argument has not been responsible for one’s acceptance of either. Diagnosis D, by point-

ing out the conditions under which the acceptance of the premises is contingent on the prior acceptance of the conclusion, shows us why circular arguments are less effective.

To conclude, when calling an argument against a skeptical hypothesis “circular”—a claim that is intended to be devastating—it is useful to provide an analysis of why this is so. The advantage of providing a diagnosis of circularity is that it helps us to distinguish between circular and non-circular arguments in cases where our intuitions do not give us much insight into the argument’s logical structure. Exploring the sources of circularity can encourage us to avoid circular anti-skeptical arguments and can show us the importance of appealing to those who can be properly convinced: those who are open to rational persuasion and who are not already persuaded.

Received February 2009

*Revised March 2009**

** My thanks to Geoffrey Pynn, of Northern Illinois University, for his very helpful comments and suggestions.*

ENDNOTES

- 1: It has recently been shown that a congenitally blind person is capable of recovering vision even after twelve years of blindness. See: Ostrovsky, Andalman, and Sinha, “Vision Following Extended Congenital Blindness,” *Association for Psychological Science Research Report* 12 (17), 1009-1014, 2006.
- 2: Robert Nozick, *Philosophical Explanations*. (Cambridge MA: Harvard University Press, 1981), 172-176.
- 3: Diagnosis D must be evaluated with respect to all the premises of an argument, because the purpose of D is to show that an argument is circular if S would still believe at least one of the premises even when this premise is false. If it can be shown that at least one of the premises fits the criteria outlined in D1 and D2, then the argument is circular under D. In order to show that D does not apply to an argument, one must show that D does not apply to any of its premises.
- 4: Nozick, 199.
- 5: C. Hurovitz, S. Dunn, G.W. Domhoff and H. Fiss, “The dreams of blind men and women: A replication and extension of previous findings,” *Dreaming* 9 (1999): 183-193.
- 6: Nozick, 156.
- 7: René Descartes, “Meditations on First Philosophy,” from *Selected Philosophical Writings*, trans. Cottingham, Stoothoff, and Murdoch, (New York: Cambridge University Press, 1988), 106.
- 8: Keith DeRose, “Descartes, Epistemic Principles, Epistemic Circularity, and Scientia,” *Pacific Philosophical Quarterly* 73 (1992): 224.
- 9: Descartes, 96.
- 10: In particular, Descartes uses the causal principle of ideas to assure himself that the cause of the idea of a perfect God—that is, God—must be at least as real as the idea itself.