Knowledge, Luck, and the Cartesian Circle

Abstract. This paper explores a widely overlooked interaction between Descartes' epistemology and his views on freedom and responsibility, with the broader aim of explaining a famously puzzling feature of his response to skepticism. In the Fourth Meditation, Descartes attempts to vindicate the reliability of reason by using his own faculty of reason, even though he appears not to have accepted the legitimacy of using sensory perception or testimony to vindicate the reliability of sensory perception or testimony. I argue that the explanation can be found only if we follow the line that Descartes draws from knowledge to responsibility, from responsibility to voluntary control, and from voluntary control to the use of one's own faculty of reason as a means to embracing the truth. The upshot is a novel form of the "simple picture" of how the Meditator comes to know the general truth that every clear and distinct perception is true, according to which this general truth is known simply in virtue of its being clearly and distinctly perceived.

1. The circle

Descartes ends the Fourth Meditation with the conclusion that everything he clearly and distinctly perceives is true. Was his procedure for arriving at this conclusion viciously circular?

In the First Meditation, Descartes considers the possibility that he was created not by God but by an original cause less perfect than God, so that Descartes' own nature is so imperfect that he is in error even about even those things which seem to him most evident. But then Descartes goes on in the Fourth Meditation to conclude that his nature is not so imperfect after all. He does so by first concluding that his original cause is, after all, a non-deceiving God. From this Descartes infers that he will never be led to error if he uses his faculty of judgment correctly, which is alleged to consist in assenting to all and only those things which he clearly and distinctly perceives. Descartes then concludes that everything he clearly and distinctly perceives is true. In the Principles, Descartes describes the significance of this result as follows:

This disposes of the most serious doubt, which arose from our ignorance about whether our nature might be such as to make us go wrong even in matters which seemed to us utterly evident (CSM I 203, AT VIIA, 16).

It is not hard to see why this procedure has struck many as viciously circular. Prior to mounting an argument that his original cause is a non-deceiving God, Descartes regards as doubtful the general trustworthiness of his own faculties. After considering the argument, however, he claims to know that his faculty for having clear and distinct perceptions is not imperfect in the way he was worried about. But it seems that in evaluating both the argument’s

---

1 I wish to thank Ned Block, Paul Boghossian, Jonathan Cottrell, Don Garrett, Colin Marshall, John Morrison, Elliot Paul, Jim Pryor, Stephen Schiffer, and the participants of the NYU Spinoza Reading Group and the NYU Dissertation Seminar.
validity and the truth of its premises, Descartes must rely on the very faculty whose general trustworthiness he hopes to verify. Even assuming that the argument is formally valid and that its premises are true, it may seem that for this reason Descartes’ belief in its conclusion could not be justified by his appeal to it, and therefore could not amount to knowledge.

In what follows, I will attribute to Descartes a view of knowledge and of justification under which his procedure can confer knowledge that everything he clearly and distinctly perceives is true. As I will argue in Sections 2 and 3, Descartes’ considered view is that all that is required for knowledge is clear and distinct perception, and so even prior to his concluding in the Fourth Meditation that his clear and distinct perceptions are true, he still knows those things which he clearly and distinctly perceives. When he goes on to conclude that all his clear and distinct perceptions are true, he takes himself to know this because he takes himself clearly and distinctly to perceive it.

The account I will advocate, then, is an instance of what Janet Broughton (2002, pp. 175-186) calls the “simple picture.” According to this picture, each particular clear and distinct perception is known to be true even before Descartes reaches the general conclusion that all clear and distinct perceptions are true. Furthermore, this general conclusion is known as soon as he clearly and distinctly perceives that it is true, and solely because he clearly and distinctly perceives its truth. In contrast, the “two-levels” picture of Feldman (1975) and DeRose (1992) says that no particular clear and distinct perceptions are known with certainty prior to the conclusion that every clear and distinct perception is true. And according to Broughton’s own interpretation, and also the interpretation presented in Newman & Nelson (1999), only a proper subset of clear and distinct perceptions are known before arriving at the general conclusion—although these interpretations differ about what subset that is.²

While advocates of these competing interpretations present textual evidence for their views, I take it that the driving motivation is not a textual but rather a philosophical dissatisfaction with the simple picture. Prima facie, it strikes many interpreters as viciously circular to appeal to one’s clear and distinct perceptions in order to prove that one’s clear and distinct perceptions are true. I will address this philosophical objection indirectly, first by outlining a general account of justification in Section 4 according to which assent to clear and distinct perception is justified, and then using this account in Section 5 to contrast Descartes’ procedure for judging that his clear and distinct perceptions are true with the superficially analogous procedure of judging that everything one’s teacher says is true because one’s teacher says so. My aim is to explain why it makes sense, according to the account I attribute to Descartes, to differentiate between this obviously vicious procedure and the one that Descartes employs.

Finally, in Section 6 I will respond to some textual objections to the simple picture presented by its opponents, the most general being that the picture robs of its point Descartes’ project of proving the truth of clear and distinct perceptions. As I interpret Descartes, there are many

² For a distinct but overlapping taxonomy of interpretations, see Hatfield (2002). The interpretation advocated in the present paper qualifies as a “strong validation” interpretation under Hatfield’s taxonomy, since it attributes to Descartes the project of offering a positive validation of the reliability of reason. Alternative interpretations see Descartes as pursuing more limited aims, such as that of establishing mere psychological certainty rather than truth. For further discussion of this “certainty, not truth” strategy, see Frankfurt (1970) and Loeb (1992).
particular truths that can be known without one’s knowing the general truth that everything one clearly and distinctly perceives is true. So I need to explain why it is nevertheless crucial on the Cartesian picture to know this general truth. In the course of doing so, I will explain Descartes’ distinction between \textit{cognitio} and \textit{scientia}, and defend my own interpretation of the former of these as sufficient for knowledge and for certainty.

2. Knowledge, justification, and rules of judgment

After concluding in the Fourth Meditation that everything he clearly and distinctly perceives is true, Descartes tells us that by doing so he has determined “not only what precautions to take to avoid ever going wrong, but also what to do to arrive at the truth” (AT VII 62, CSM II 43). The precautions for avoiding error involve following this rule:

\begin{quote}
\textbf{Error-Avoidance Rule:} If you do not clearly and distinctly perceive that P, then don’t judge that P!
\end{quote}

And what to do to arrive at the truth is to follow the converse rule:

\begin{quote}
\textbf{Truth Rule:} If you clearly and distinctly perceive that P, then judge that P!
\end{quote}

In a sense, these rules are nothing new. In the Fourth Meditation, Descartes has made some important discoveries about the nature of his own mental faculties. Apparently, their nature is such that he is able to withhold his assent from any perception which is not clear and distinct. So it seems that his following the error-avoidance rule was the necessary result of his initial resolution to doubt everything he was able to doubt. Conversely, Descartes has discovered that he in some sense is unable to withhold his assent from what he clearly and distinctly perceives. So it seems that following the truth rule is something Descartes also must have been doing all along, even if he was not always aware that this is what he was doing at the time.

Descartes tells us so in the Third Meditation, when he reflects on the \textit{cogito} argument of the preceding Meditation:

\begin{quote}
I am certain that I am a thinking thing. Do I not therefore also know what is required for my being certain about anything? In this first item of knowledge there is simply a clear and distinct perception of what I am asserting; this would not be enough to make me certain of the truth of the matter if it could ever turn out that something which I perceived with such clarity and distinctness was false. So now I seem to be able to lay it down as a general rule
\end{quote}

\footnote{Here is Descartes on how it is possible to make discoveries about the nature of one’s mental faculties: As to the fact that there can be nothing in the mind, in so far as it is a thinking thing, of which it is not aware, this seems to me to be self-evident. … But it must be noted that, although we are actually aware of the acts or operations of our minds, we are not always aware of the mind’s faculties or powers, except potentially. By this I mean that when we concentrate on employing one of our faculties, then immediately, if the faculty in question resides in our mind, we become actually aware of it, and hence we may deny that it is in the mind if we are not capable of becoming aware of it. (AT 246-247, CSM II 171-172)
that whatever I perceive very clearly and distinctly is true. (AT VII 35, CSM II, 24)

This passage, as well as the several paragraphs which follow it, are notoriously difficult to interpret. Although it is not clear why, Descartes is at this stage hesitant in his assertion that everything he clearly and distinctly perceives is true, and he seems to find it necessary to give a much longer argument for the same conclusion in the central thread of the Third and Fourth Meditations. But whatever explains this hesitance concerning the general conclusion about the truth of clear and distinct perceptions, Descartes is less hesitant when it comes to the following more particular claims: First, he is certain that he is a thinking thing. Second, this qualifies as an item of knowledge. And third, in this item of knowledge there is nothing but a clear and distinct perception of what he knows.

However we ought to understand the hesitance of this passage, there can be no question that Descartes’ later, considered view was indeed that every clear and distinct perception is true. On the interpretation I wish to advance, it was furthermore his considered view that all that is required for knowledge is clear and distinct perception. This further claim is crucial to our understanding of the Cartesian Circle, for it implies that a higher-order vindication of the epistemic rules one follows is not necessary for first-order knowledge.

To get a better understanding of this claim, and to address the apparent problems it faces, we can begin with some preliminary points about the notion of a rule—a notion which plays a central role in all of Descartes’ major philosophical works.

A rule tells you to perform a given action. A conditional rule tells you to perform an action provided that certain antecedent conditions are met. These rules take the form: If C, then do A! When the relevant act is one of granting or withholding judgment, we can say that the rule in question is a rule of judgment. Rules of judgment will be of particular interest for understanding Descartes’ conception of knowledge because in Descartes’ view, only a judgment can qualify as an instance of knowledge or of error.

The truth rule is a rule of judgment whose antecedent condition C is that one clearly and distinctly perceives what one judges to be true. Conversely, the error-avoidance rule says not to

---

4 This seems to reflect a change in his thinking since the publication of the Discourse, in which an argument like the one from the Third Meditation seems to take priority. The premise that our clear and distinct perceptions are true, in conjunction with the further premise that this will be assured only if God exists, are there offered as support for the existence of God. See Van Cleve (1979) for discussion.

5 It may seem that Descartes withdraws this claim two paragraphs later, where he says “[I]f I do not know that there is a God, then it seems that I can never be quite certain about anything else” (AT VII 36, CSM II 25). But when Descartes is pressed in the Second Objections on whether this means that he was not in fact certain that he was a thinking thing until he proved that God exists, Descartes responds:

When I said that we can know nothing for certain unless we are aware that God exists, I was expressly declared that I was speaking only of those conclusions which can be recalled when we are no longer attending to the arguments by means of which we deduced them. (AT VII 140, CSM II 100)

See Section 6 for more on the distinction between cognitio and scientia.
make a judgment when this antecedent condition fails to obtain. These two rules occupy a privileged position in Descartes' epistemology.\(^6\)

We can say that a rule of judgment is knowledge-conducive just in case, whenever you arrive at a true judgment by following the rule, that judgment qualifies as knowledge. What conditions will a rule of judgment have to meet in order to be knowledge-conducive?

First, it is plausible that a knowledge-conducive rule must be truth-conducive. Since Descartes took knowledge to require infallibility, the notion of truth-conduciveness of greatest relevance to his view is correspondingly an infallibilist one. On this infallibilist understanding, a rule is truth-conducive only if one’s following the rule guarantees that one’s resulting judgment is true. For our purposes, it will do no harm to follow Descartes in understanding truth-conduciveness in this infallibilist way, for it will do no harm to set the bar for knowledge implausibly high.

Of any particular rule \( R \), one might wonder whether \( R \) is truth-conducive. In addition to this, an unreflective thinker, who does not know what rule (or rules) of judgment he follows, can ask whether the rule(s) he follows, whatever they are, are truth-conducive. We can see the Descartes of the First Meditation as asking this, in response to his worries about his own imperfection. As I will explain shortly, we should see the Descartes of the Fourth Meditation as answering this question by claiming that he has been following the truth rule, and by concluding that this rule is truth-conducive.

We are now in a position to clarify the charge of the critic who says that Descartes is guilty of vicious circularity. According to the critic, Descartes' procedure for concluding that that the truth rule is truth-conducive is viciously circular, and cannot confer knowledge. To be clear, the charge is not that Descartes' conclusion is false, but rather that Descartes' procedure cannot confer knowledge of its truth even if it is true. Since Descartes' procedure for arriving at this conclusion consists in following the truth rule, the critic must deny that the truth rule is knowledge-conducive. And since Descartes' conclusion is that the truth rule is truth-conducive, the critic must furthermore say that the truth rule fails to be knowledge-conducive even if it truth-conducive. We can therefore grant to Descartes that the truth rule is truth-conducive, and ask what further condition for knowledge-conduciveness it might fail to meet.

Assuming that justification is necessary for knowledge, another necessary condition for a rule to be knowledge-conducive is that it be justification-conducive. That is, the rule must be such that whenever one follows it, one's resulting judgment is justified. Presumably, the critic who charges Descartes with vicious circularity wishes to deny that the truth rule meets this condition. The rough idea, I take it, is that if Descartes has general doubts about his clear and

\(^6\) In the Discourse and the Principles, Descartes lays down a number of rules for "philosophizing correctly" which are not rules of judgment. For example, he gives as rules: Lay aside your preconceived ideas!, and Divide questions into as many parts as possible!. These rules, unlike the truth rule, do not lead immediately to knowledge. A judgment can constitute knowledge, but the act of laying aside preconceived ideas, or of dividing a question into many parts, cannot. These other rules, although not rules of judgment, do contribute to knowledge indirectly, since they help you to make your ideas clear and distinct. That is, they help to bring about the antecedent conditions for making a judgment which would then amount to knowledge.
distinct perceptions, then he could not be justified in following a rule which instructs him to 
assent to whatever he clearly and distinctly perceives.

Was it Descartes' considered view that the truth rule is justification-conducive? That is, did he 
think that one is justified whenever one makes a judgment by following the truth rule? In my 
view he did think this, although it will take some work to show it.

Suppose we grant that Descartes' considered view was in agreement with the hesitant 
assertions of the above Third Meditation passage, in which he says that all which is required to 
make him certain is a clear and distinct perception. It is not obvious that this implies that he 
always is justified in making judgments by following the truth rule, for at least two reasons.

First, following a rule cannot be a matter merely of conforming to the rule's instructions. If 
one judges that P by following the truth rule, then in some sense one's reason for judging that P 
must be one's clear and distinct perception that P. The distinction between following a rule and 
merely conforming to it makes room for a view like Don Garrett's (2004), which says that 
Descartes' judgments in the early Meditations are supposed to be justified by their indubitability. 
On this view, the clarity and distinctness of Descartes' perceptions causally explain their indubitability, but it is indubitability, rather than their clarity and distinctness, which serves to 
justify Descartes' judgments.

Second, Descartes had a general tendency to run together factual questions about how the mind 
in fact makes judgments with normative questions of how it ought to. This is perhaps 
understandable, since Descartes thought the mind had as an original cause a non-deceiving God. 
Nevertheless, the conflation presents us with obstacles in trying to interpret him. For example, 
Jonathan Bennett (1990) has argued persuasively that when Descartes says that that P is 
indubitable, he sometimes is most naturally read as meaning that there is no good reason to 
doubt that P, and at other times simply as meaning that he is psychologically incapable of 
doubting that P. Sometimes, it is not clear what he means.

Although Descartes did not always distinguish the descriptive claim that clear and distinct 
perceptions cause one to be certain from the normative claim that they justify certainty, I think 
it is clear that he accepted both. In the Meditations and the Principles, Descartes speaks 
frequently of judgments being “correct” or “incorrect,” and I think we can understand these 
terms to mean something roughly along the lines of what is today meant by “justified” and 
“unjustified.” And what Descartes says is that the correct use of judgment is a matter of 
assenting to what you clearly and distinctly perceive, rather than of assenting to what is 
indubitable. The indubitability of clear and distinct perceptions appears only as part of an 
argument that assenting to them is correct. This is the view I will attribute to him in what 
follows.

3. Justification and the correct use of judgment

We can begin with a closer look at a portion of Descartes' Fourth Meditation argument that 
everything he clearly and distinctly perceives is true. After claiming that he has within him a 
faculty of judgment whose original cause is a non-deceiving God, Descartes concludes:

(1) My faculty of judgment is one which could never lead me to go wrong when
I use it correctly (recte/comme il faut).

Descartes then presents an argument, which we will discuss below, for the claim that

(2) If I assent to a clear and distinct perception, then I am using my faculty of judgment correctly.

From (1) and (2) it follows that when he assents to a clear and distinct perception, he will never “go wrong,” and therefore that:

(3) Everything I clearly and distinctly perceive is true.

We see here that (2), a crucial premise in Descartes’ argument for (3), appeals to a claim about the circumstances in which he will be using his faculty of judgment correctly. And what this premise says is that he will be using his judgment correctly if he assents to what is clearly and distinctly perceived.

It is important that (2) is a premise in Descartes’ argument for (3). Descartes cannot be saying simply that once he knows (3), he will be using his faculty of judgment correctly whenever he assents to a clear and distinct perception. Rather, the argument goes the other way around. He argues for (3) by appealing to the premise that assenting to clear and distinct perceptions is “correct,” a premise which is known and therefore true even before he knows that everything he clearly and distinctly perceives is true.

On the interpretation I wish to press, Descartes’ considered view is that whenever he assents to a clear and distinct perception, his assent is justified, and that it is justified because the perception assented to is clear and distinct (i.e., not because it is indubitable). So what I now want to argue is, first, that (2) should be understood as claiming that his assent is correct because the perception assented to is clear and distinct, and second, that when one’s assenting to a perception is correct, it is justified in the sense required for knowledge.

In the Fourth Meditation, Descartes’ strategy for determining what constitutes a correct use of judgment is first to determine what constitutes an incorrect use, or misuse, of judgment. And so, most of Descartes’ explicit statements about how to use one’s judgment concern its incorrect uses. And so we find him making claims like the following:

…it surely is no imperfection in God that he gave me the freedom to assent or not to assent in those cases where he did not endow my intellect with a clear and distinct perception; but it is undoubtedly an imperfection in me to misuse that freedom and make judgments about matters which I do not fully understand. (AT VII 61, CSM II 42)

A few paragraphs earlier, in what I will call the “Essence of Error” passage, he begins by telling us how to avoid the misuse of judgment:

If … I simply refrain from making a judgment in cases where I do not perceive the truth with sufficient clarity and distinctness, then it is clear that I am behaving correctly.
and avoiding error. But if in such cases I either affirm or deny, then I am not using my free will correctly. (AT VII 59-60, CSM II 41)

In these passages, Descartes tells us that it is a misuse of his freedom when he makes judgments on matters which his intellect does not clearly and distinctly perceive. But it seems that he also must have thought that assenting to a clear and distinct perception constitutes a correct use of judgment, since his argument for (3) depends upon the claim that assenting to clear and distinct perceptions is correct.

This is further confirmed by a passage in the Second Objections. When pressed on whether everything he clearly and distinctly perceives is true, he replies that his original cause is a non-deceiving God, adding:

Hence this faculty [of judgment] must tend towards the truth, at least when we use it correctly (that is, by assenting only to what we clearly and distinctly perceive, for no other correct method of employing this faculty can be imagined). (AT VII 144, CSM II 103)

So in other words, when you assent to a clear and distinct perception, your assent is correct. And it is correct because what you assent to is clearly and distinctly perceived.

Turn now to my second claim, which is that when one’s assent is correct, it is justified. Let’s start by noting some properties commonly associated with justification.

First of all, it is sometimes claimed that justification must be, at least under ordinary conditions, somewhat conducive to truth. There is no question that Descartes thought the correct use of judgment was conducive to the truth.

Second, justification is often said to be necessary for knowledge. As we will see below, Descartes in the Essence of Error passage says that the correct use of judgment is necessary for knowledge.

Third, and most importantly, justification is a normative status. Although this is controversial, it is plausible to understand this normative status deontologically—roughly speaking, in a way associated with responsibility, praise, and blame. As we will see, Descartes seemed to think that the correct use of judgment is praiseworthy, and that its misuse is blameworthy.

Now Descartes’ argument for (3) does not require that correctness be a positive normative status of judgment. Rather, it requires only that the correct use of judgment must in some sense be in keeping with a non-deceiving God’s intentions for its use. But I think Descartes took the correctness or incorrectness of a judgment to have normative implications.

After telling us in the Essence of Error passage how to avoid the misuse of judgment, Descartes goes on to say this:

If [in a case where I lack clear and distinct perception] I go for the alternative which is false, then obviously I shall be in error; if I take the other side, then it

---

7 For further discussion, see, e.g., Alston (1988), Plantinga (1988), and Weatherson (2008).
is by pure chance that I arrive at the truth, and I shall still be at fault since it is clear by the natural light that the perception of the intellect should always precede the determination of the will. In this incorrect use of free will may be found the privation which constitutes the essence of error. (AT VII 60, CSM II 41)

Here, Descartes says that in the absence of clear and distinct perception, even if he assents to the truth, he will not have avoided being “at fault” (culpa / fallir). This is clearly a negative normative evaluation of his assent.

In the *Principles* (38), he considers the “defect in the way we act” which results in error, and compares it to a sin. And in the Fourth Meditation, he says that the indifference in the will, which is present in cases where clear understanding is lacking, is responsible for both error and sin (AT VII 58, CSM II 40-41). He takes a similar line in a passage from the *Second Replies* as well:

If, despite the fact that [Christian] doctrines are obscure to [an atheist], he is induced to embrace them by fallacious arguments, I make bold to assert that he will not on that account be a true believer, but will instead be committing a sin by not using his reason correctly (AT VII 148, CSM II 106).

Here Descartes says that judgments made on the basis of fallacious arguments are sinful, even when what one assents to is true. Notice also that this is sinful not because the nonbeliever is supposed to know that his embracing the conclusions of a fallacious argument constitutes an incorrect use of his judgment, but instead merely because he is, as a matter of fact, “not using his reason correctly.”

So Descartes thought that one is “at fault,” and apparently even “committing a sin” when one grants assent through an incorrect use of one’s judgment—even if what one assents to is true. But what does he mean when he says that in this can be found the “privation” which constitutes the essence of error? I take it that he has in mind the privation of knowledge which constitutes the essence of error. This is suggested by a passage earlier in the Fourth Meditation, where he says that his errors, as opposed to his mere ignorance of truths, constitute “a privation or lack of some knowledge which somehow should be in me” (AT VII 55, CSM II 38, emphasis added).

This interpretation is further reinforced in a corresponding passage in the *Principles*, where Descartes says:

44. *When we give our assent to something which is not clearly perceived, this is always a misuse of our judgment, even if by chance we stumble on the truth.*

He explains the significance of this point as follows:

If we do stumble on the truth, it is merely by accident, so that we cannot be sure that we are not in error. (AT VIIIa 21, CSM I 207)

---

8 I am here assuming that Descartes believed that Christian doctrines are true, or at least that he wished to present himself as believing that they are true.
If I cannot be sure that my judgment that P is not in error, then I cannot be sure that P. So Descartes is saying here that I cannot be sure that P if my assent to P constitutes a “misuse” of my faculty of judgment. I think the idea here is that being sure that P, like knowing that P, requires the correct use of judgment.

What have we learned? First, the correct use of judgment is conducive to truth. Second, if you assent to the truth by misusing your faculty of judgment, you will be at fault. And third, when your assent constitutes a misuse of judgment, you won’t be sure of what you judge, and you will be deprived of knowledge. Given all of this, it seems reasonable to attribute to Descartes the view that what we nowadays call justification consists in the correct use of judgment. And the correct use of judgment, as we have seen, consists of assenting to all and only what you clearly and distinctly perceive. In short, we have learned that Descartes thought the truth rule is justification-conducive.

4. Rules for the direction of the mind

Even if it is true that one’s assent to clear and distinct perceptions is justified, and that it amounts to knowledge, this should not simply be a brute fact about the nature of knowledge and of justification. Rather, we ought to be able to say more generally what it means to be justified in judging that P, and for one’s judgment to amount to knowledge, in a way which explains why judgments arrived at by following the truth rule are justified and amount to knowledge. Descartes himself never addresses the issue at this level of abstraction—his goal is simply to determine what rule or rules are knowledge-conducive and then to follow them, rather than to give a general theory of knowledge which explains why they are knowledge-conducive. But I do think that he gives us enough clues about how he thought of the nature of knowledge (implicitly, at least) that we can piece together the rough outlines of a general account.

When first introducing his rules for making judgments in the Discourse, Descartes tells us that it was his dissatisfaction with his teacher’s opinions which left him “forced to become [his] own guide”—an enterprise which he says “has never gone beyond trying to reform [his] own thoughts and construct them on a foundation all [his] own” (AT VI 15-16, CSM I 118-119). The ultimate goal of this enterprise of course is to arrive at beliefs which are true, but it is crucial for Descartes that his means for pursuing this goal involve the use of his own reason, or faculty of judgment. “Since God has given each of us a light to distinguish truth from falsehood,” he goes on to tell us in Part Three, “I should not have thought myself obliged to rest content with the opinions of others for a single moment if I had not intended in due course to examine each of them using my own judgment” (AT VI 27, CSM I 124).

It is perhaps appropriate, then, that the rules he goes on to formulate are what might be called individualistic rules. That is, they are rules for guiding a single individual in his own solitary

---

9 The cited passages could be taken to mean only that the correct use of judgment is necessary for justification. But as we have seen, Descartes also says that clear and distinct perception is sufficient for knowledge (which requires justification). He furthermore seems to imply in the Fourth Meditation that the explanation of his inability to withhold assent from clear and distinct perceptions is that he “clearly understands that the reasons of truth … point that way” (AT VII 58, CSM II 40, emphasis added). In any case, it seems unnatural to read Descartes here as saying that the correct use of judgment is necessary for justification, but that some further (unspecified) condition also is necessary.
search for the truth—not for an enterprise of collective research and instruction. As I interpret him, Descartes thought justification, and hence knowledge, can be achieved only through the use of individualistic rules or methods. After presenting the motivation for this view in broad outline, I will in Section 5 discuss its relevance to the Cartesian Circle.

Descartes gives a number of rationales for this individualism in the Discourse and elsewhere, but here I will focus on what I take to be the best rationale available to him. As we will see, this rationale for using individualistic rules leads naturally to the use of purely mental rules—rules which concern solely the internal states of one’s own mind—both in terms of what they instruct you to do (e.g., make a judgment), and in the antecedent conditions under which they tell you to do it (e.g., when you are unable to withhold assent). In fact, while Descartes in his philosophical works formulates many rules which he says should guide us in our search for the truth, all, or almost all, are internal rules in this sense.10 Indeed, Descartes has only two works which mention rules (regulae) or methods (methode) in their titles, these being: Rules for the Direction of the Mind (Ingenii), and the Discourse on the Method of rightly conducting one’s reason and seeking the truth in the sciences. The key to understanding Descartes preoccupation with mental rules will be the line that he draws from knowledge to praiseworthiness, from praiseworthiness to voluntary control, and from voluntary control to the use of mental rules. I will start at the end of this line and work backwards.

After introducing his rules for the direction of his reason in the Discourse, Descartes tells us of the rules he followed provisionally in his practical affairs. One of these was to “become accustomed to believing that nothing lies entirely within our power except our thoughts, so that after doing our best in dealing with things external to us, whatever we fail to achieve is absolutely impossible so far as we are concerned.” The result of our making this resolution is that “we shall not regret the absence of goods which seem to be our birthright when we are deprived of them through no fault of our own.” The result, he says, is that we will be freer and happier than others who lack this resolution, “however favored by nature and fortune they may be” (AT VI 25-26, CSM I 123-124).11

This rather extreme passage has two central themes. The first is that, insofar as the external world is concerned, we are the potential victims of fortune and chance. In contrast, within our own minds, our control is potentially absolute. The second central theme of the extreme passage from the Discourse is that we can escape from the dominion of fortune by valuing only what is internal to the mind.

---

10 There are a few rules in the Rules for the Direction of the Mind which instruct you to use written symbols and diagrams in order to aid your imagination and memory. Still, the use of written symbols is just supposed to play the heuristic role of helping you to better use your memory and imagination.

11 Although in the Discourse Descartes says only that he has accepted this belief provisionally, he retained views similar to these in his later years. In a letter to Elizabeth in which he cites the above passage from the Discourse, Descartes says:

It seems to me that each person can make himself content by himself without any external assistance, provided he respects three conditions … The third is that he should bear in mind that while he thus guides himself as far as he can by reason, all the good things which he does not possess are one and all entirely outside his power. (AT IV 265, CSMK III 257-258)
Descartes gives a lucid statement of a *prima facie* similar view in *The Passions of the Soul*:

I see only one thing in us which could give us good reason for esteeming ourselves, namely, the exercise of our free will and the control we have over our volitions. For we can reasonably be praised and blamed only for actions that depend upon this free will. (AT XI 445, CSM I 384).

Despite the similarities, there is a crucial difference in what these passages suggest. The more extreme passage from the *Discourse* suggests that all of our evaluative attitudes should be restricted to what is internal to the mind. The less extreme passage from the *Passions*, in contrast, speaks only of esteem, praise, and blame. These evaluative attitudes are the ones normally associated with deontological notions like responsibility and justification, so for our purposes all we need to observe is that Descartes held the less extreme doctrine. We can be praised and blamed, then, only for what is under our free and voluntary control. If our control is limited only to what is internal to the mind, as Descartes suggests, then it follows that we can be praised or blamed only for what is internal to the mind.

Returning now to Descartes' discussion of the truth rule, we find in the *Principles* that:

37. *The supreme perfection of man is that he acts freely or voluntarily, and it is this which makes him deserve praise or blame.*

...[man's freedom] makes him in a special way the author of his actions and deserving of praise for what he does. We do not praise automatons for accurately producing the movements they were designed to perform, because the production of these movements occurs necessarily. It is the designer who is praised for constructing such carefully-made devices ... By the same principle, when we embrace the truth, our doing so voluntarily is much more to our credit than would be the case if we could not do otherwise. (AT VIIIa 18-19, CSM I 205)

Here Descartes says that since only what is voluntary can be praiseworthy, when we assent to the truth our doing so is praiseworthy only if it is under our voluntary control. This is ambiguous between a weaker and a stronger interpretation. The weaker interpretation says that we can be praiseworthy for our assenting to the truth only if *assenting* is within our control, while the stronger says that assenting to the truth can be praiseworthy only if *assenting to the truth* is under our voluntary control. For comparison, giving money to UNICEF is praiseworthy, while giving money to the KKK is blameworthy. But in order for you to be praiseworthy for giving money to UNICEF, it is not enough that you have control over whether you give money away to an organization. Rather, you must have control over whether the organization you give to is UNICEF rather than the KKK.

I think the stronger of the two interpretations is probably the better reading of the passage, and in any case it makes better sense of Descartes' reliance on mental rules if we suppose that he accepted it. For if we assume as Descartes did that you have control only over what is internal to the mind, then your assenting to the truth can be under your control only if it can be
accomplished by means which are internal to the mind. And for this reason, your assent can only be praiseworthy, or justified, if it is accomplished by the use of mental rules.

In a bit more detail, the idea here is this. Your assenting to the truth was under your control only if you followed a rule such that: 1) your following the rule is under your control, and 2) following the rule leads without exception to true judgments (i.e., the rule is truth-conducive). If you assent to a truth by following a rule meeting both these conditions, then your assenting to the truth was under your voluntary control, and so you are praiseworthy for having done so. And if praiseworthy assent to the truth (i.e., justified true belief) is sufficient for knowledge, then your judgment will amount to knowledge.\footnote{It should be noted that, even though Descartes believed assent is free and voluntary, he took this to be consistent with clear and distinct perceptions being indubitable. For example, in the Fourth Meditation he says this:

[The will, or freedom of choice] consists simply in the fact that when the intellect puts something forward for affirmation or denial or for pursuit or avoidance, our inclinations are such that we do not feel we are determined by any external force. In order to be free, there is no need for me to be inclined in both ways; on the contrary, the more I incline in one direction ... the freer is my choice. (AT VII 57-58, CSM II 40)

For further discussion, see Cottingham (2002) and Weatherson (2008).}

5. Reliability and justification

If Descartes is right that judgments arrived at by the truth rule amount to knowledge, it is very easy to see that his procedure for assenting to his conclusion that his clear and distinct perceptions are true could not be circular in a way which rules out its being known. For if he is right that the premises of his argument are clearly and distinctly perceived, and also that it is clearly and distinctly perceived that his conclusion follows from these premises, then his judging that his clear and distinct perceptions are true will simply be an instance of the truth rule.

Of course, this shows merely that a particular account of knowledge and of justification has as a consequence that Descartes' judgment that his clear and distinct perceptions are true is epistemically justified by his procedure. This may seem to say more about the plausibility of this account of justification than it does about the justification of Descartes' judgment. For comparison, if you doubted the general veracity of an epistemic authority figure—such as a teacher or a religious text—it would be viciously circular to respond to this doubt by judging that everything the authority says is true simply on the basis of the authority's saying so. Any account of justification according to which such a judgment would be justified should for that reason be rejected. So what, on the Cartesian account, makes Descartes' procedure for judging that his clear and distinct perceptions are true any different?

To answer this, we should first get clearer about why a judgment that everything the authority says is true, based solely on the authority's saying so, is not justified. In an important sense, the problem with this judgment is not circularity. Rather, circularity is simply an instance of a more general problem.

Suppose you have wholly general doubts about the truth of what your teacher says. If so, you would be unjustified in believing that everything your teacher says is true simply because your
teacher says so. But you also would be unjustified in believing other claims your teacher makes simply because he says they are true. This is because you doubt the trustworthiness of your teacher in general—it is not that you think your teacher is particularly disposed to error on the matter of whether everything he says is true.

Correspondingly, when Descartes worries in the First Meditation that his own faculty of clear and distinct perceptions is imperfect, he is not worried that he is particularly disposed to error when it comes to the question of whether his faculty of clear and distinct perception is imperfect. Rather, the worry is that he is prone to error in general. The upshot is that if he is unjustified in accepting that clear and distinct perceptions are true by using his own faculty of clear and distinct perception, this is just because he would be unjustified in assenting to anything by using his faculty of clear and distinct perception.

So what we want to know is what the difference is supposed to be between judging that \( P \) because you clearly and distinctly perceive that \( P \), on one hand, and judging that \( P \) because your teacher says so, on the other, such that the former judgment is justified even in the absence of independent reason for thinking that your clear and distinct perceptions are true. If we could answer this in the case of some other proposition which Descartes thought we clearly and distinctly perceive, it is hard to see what additional challenges are presented by the special case where \( P \) is the proposition that your clear and distinct perceptions are true.

Now we concluded Section 4 by saying that in order for you to be able to know that \( P \), you must be able to arrive at a judgment that \( P \) by a generally truth-conducive rule, the following of which is under your voluntary control. Supposing you know that your teacher says that \( P \), is there a rule meeting these conditions which would instruct you to assent to this? An obvious suggestion is this:

\begin{quote}
**Rule A:** If you know that your teacher says that \( P \), then judge that \( P \)!
\end{quote}

But this rule is not truth-conducive. Teachers often say things that are false, and so it will not be true in general that following Rule A leads to true judgments. It may be that your teacher happens to say something true on a given occasion, or even that you happen to have a teacher who says only things that are true. But because it is not true in general that Rule A leads to true judgments, your assenting to what he says by this rule would not enable you to control whether you were assenting to the truth. And this means that you cannot be praiseworthy, or justified, for having assented to the truth.

We might try to build into the antecedent conditions of our rule a guarantee that your teacher says something true—for example, by requiring that you have a teacher who only speaks the truth. That is:

\begin{quote}
**Rule B:** If you know that your teacher says that \( P \) and if in fact everything your teacher says is true, then judge that \( P \)!
\end{quote}

Like the truth rule, this rule will be truth-conducive. But unlike the truth rule, whether you follow this rule will not in general be under your control. If you try to follow it without
knowing whether everything your teacher says is true, whether you succeed in following the
rule will not be under your control. And so even if you end up following it, and therefore end
up assenting to a truth, you will not be praiseworthy for assenting to the truth because your
following the rule will not have been under your voluntary control.
Finally, there is this rule:

Rule C: If you know that your teacher says that \( P \) and if you know everything your
teacher says is true, then judge that \( P \)!

This rule is fine by Descartes’ standards. And supposing (as Descartes would not have) that you
can know what your teacher says, and also that everything your teacher says is true, without
knowing that God exists, it could potentially confer knowledge that God exists. But this rule
will not be able to confer knowledge that everything your teacher says is true, since the
conditions for applying it require antecedent knowledge that everything your teacher says is
true.

We have said why, on the general picture of knowledge and of justification I attribute to
Descartes, it makes sense to differentiate between Descartes’ procedure of judging that your
clear and distinct perceptions are true because you clearly and distinctly perceive it, and the
procedure of judging that everything your teacher says is true because your teacher says so. To
the extent that the appearance of vicious circularity in Descartes’ procedure is motivated by an
analogy between these procedures, we have gone some way towards making Descartes’
procedure appear philosophically acceptable.

6. The purpose of vindicating reason

I have attributed to Descartes a network of views concerning the nature of knowledge and of
justification, according to which you can be justified in judging that \( P \), and know that \( P \),
whenever you clearly and distinctly perceive that \( P \). As I interpret him, Descartes thinks this is
the case even when you doubt the general truth that everything you clearly and distinctly
perceive is true. And since Descartes most likely believes that knowledge requires certainty, this
also means that you can be certain that \( P \) whenever you clearly and distinctly perceive that \( P \)—
even if you doubt the truth of your clear and distinct perceptions more generally.

According to this interpretation, one can achieve certainty of many particular truths without
being certain of the general truth that everything one clearly and distinctly perceives is true.
This view may seem unsatisfying, since it may seem to rob Descartes’ project of demonstrating
this general truth of its point. Descartes did not regard this project as a mere curiosity, as if
the general truth that everything one clearly and distinctly perceives as simply one interesting
truth which one might undertake to discover. Rather, he thought that certainty of this general
truth is crucial for a satisfying resolution of skeptical doubt, and for the development of a
science which is stable and likely to last.

To understand why this is consistent with the view I attribute to Descartes, we must bear in
mind Descartes’ repeated insistence that a clear and distinct perception justifies a judgment only
so long as that clear and distinct perception is present to the mind. Clear and distinct
perceptions which one had yesterday do not justify one’s judgments today, any more than the
proofs known by those who first introduced the proposition that there is a God can justify one's own belief.\textsuperscript{13}

Now if one knows that one had a clear and distinct perception that P yesterday,\textsuperscript{14} this knowledge can justify one in judging that P today if one also knows that everything one clearly and distinctly perceives is true—in something like the way that knowing your teacher says that P can justify a judgment that P if you know that everything your teacher says is true. But only at the moment one has a clear and distinct perception that P can the perception itself—rather than one's knowledge that there is or was such a perception, supplemented with knowledge of the general truth that clear and distinct perceptions are true—justify one's judgment that P.

For Descartes, knowledge of the general truth of clear and distinct perceptions is in this way necessary to make one justified at one time in assenting to what one knows was clearly and distinctly perceived at an earlier time. Without knowledge of this general truth, you are stuck with whatever you clearly and distinctly perceive only in the present moment. I wonder if a failure to appreciate how prevalent Descartes thought this situation is has led many commentators to mistakenly ascribe to Descartes the view that clear and distinct perception does not always justify certainty.

For example, a number of commentators claim that Descartes thought that a clear and distinct perception that $2+3=5$ fails to justify certainty that $2+3=5$, unless one knows that all clear and distinct perceptions are true. This claim is said to be supported by Descartes' discussion of the evil demon hypothesis in the First Meditation, where he says that this hypothesis brings into doubt his beliefs even in the most simple and general things, such as that $2+3=5$. Since he presumably is capable of clearly and distinctly perceiving that $2+3=5$, this has been taken to mean that $2+3=5$ is doubtful even while one clearly and distinctly perceives it (Broughton, 2002, p. 180), (Rickless, 2005, pp. 332-333, fn. 5), (DeRose, 1992).

Descartes never does say that $2+3=5$ ever is doubtful when being clearly and distinctly perceived, however. And in the Seventh Replies, he says outright that he was supposing that he was not clearly and distinctly perceiving that $2+3=5$ when he doubts it in the First Meditation:

\begin{quote}
I have explained, in several places, the sense in which [my claim that there is “nothing” which cannot be doubted] is to be understood. It is this. So long as we attend to the truth which we perceive very clearly, we cannot doubt it. But when, as often happens, we are not attending to any truth in this way, then even though we perceived many things very clearly, nevertheless there will be nothing which we may not justly doubt so long that we do not know that whatever we clearly and distinctly perceive is true. But my careful critic here takes ‘nothing’ quite differently. From the fact that at one point I said that there was nothing that we might not doubt—namely, in the First Meditation, in
\end{quote}

\textsuperscript{13} This example is discussed in The Search for Truth, (AT X 503-504, CSM II 404).

\textsuperscript{14} To know this, you would somehow have to know that your memory was not playing tricks on you. Descartes never says how this is possible, and it is a persistent and I believe insurmountable problem for his view. For further discussion, see Gewirth (1941), Doney (1955), Frankfurt (1962), and Frankfurt (1970).
which I was supposing that I was not attending to anything that I clearly perceived—he draws the conclusion that I am unable to know anything certain, even in the following Meditations. This is to suggest that the reasons which may from time to time give us cause to doubt something are not legitimate or sound unless they prove that the same thing must be permanently in doubt. (AT VII 460, CSM II 309)

Another difficulty comes from Descartes insistence that all scientific knowledge depends on knowledge of God's existence. For example, in his discussions of an atheist mathematician who clearly and distinctly perceives the proof for a theorem, Descartes says that while the atheist's belief amounts to knowledge (cognitio, or scire), it does not amount to scientific knowledge (scientia), or “perfect knowledge” (perfectio scire).

Descartes explains that it is improper to attribute scientific knowledge to the atheist because general doubts about his clear and distinct perceptions could in the future dislodge the atheist’s belief, when he is no longer attending to the proof of his theorem. However, Descartes says that the atheist will be certain (certus) of his theorem, so long as he attends to a clear and distinct perception of its proof.

Some commentators have suggested that the atheist’s certainty is meant to be only psychological certainty—an irresistible compulsion to believe the theorem. According to this view, it is only after the existence of a non-deceptive God is proved, and from this the truth of all clear and distinct perceptions, that one can arrive at normative certainty of the theorem—certainty which is justified by conclusive evidence or reasons. This “two-level” interpretation is not committed to the extreme claim that belief in the theorem has nothing going for it epistemically speaking until after the truth of clear and distinct perceptions are proved, although some of the interpretation's proponents do suggest this as well (Feldman, 1975). Rather, the two-level theorist’s core commitment is merely that the reasons for believing the theorem are epistemically inconclusive—that, while perhaps compelling, these reasons fail to justify complete certainty in the theorem.

In contrast with the two-level theorist, I think Descartes is best read as saying that the atheist mathematician is normatively certain of the theorem so long as he clearly and distinctly perceives its proof. It is only afterwards, when he no longer attends to the proof, that general doubts about clear and distinct perceptions will give him reason to doubt the theorem. The two-level theorist, who denies normative certainty to the atheist even “in the moment,” faces at least three problems.

An initial problem with the two-level interpretation is philosophical rather than textual. Descartes hopes to achieve normative certainty in his conclusion that everything he clearly and distinctly perceives is true, on the basis of premises which he clearly and distinctly perceives. If these premises have only psychological certainty, it is difficult to see how they could provide normative certainty in Descartes’ conclusion. This issue has been discussed in detail elsewhere,15 and I will not attempt to adjudicate it here.

15 See, e.g., Broughton (2002), Van Cleve (1979), and DeRose (1992).
In addition to this philosophical problem, there are textual problems facing an interpretation under which the atheist mathematician’s certainty is not justified. Although this interpretation is consistent with Descartes’ Fifth Meditation discussion of the atheist, it is rather unnatural to interpret other remarks in this way. For example, in the passage from the Seventh Replies quoted above, Descartes says that there is nothing we cannot “justly doubt,” adding that he means this to apply only when we are not attending to any truths clearly and distinctly. The implication is that it would not only be psychologically impossible, but moreover unjust to doubt things which we currently clearly and distinctly perceive. Many examples of similar statements are discussed in Della Rocca (2005).

A final source of problems concerns the psychological impossibility of lacking certainty in a clear and distinct perception. When the atheist clearly and distinctly perceives his theorem by entertaining its demonstration, the atheist is unable to refrain from being psychologically certain. According to the two-levels theorist, this certainty is unjustified. It is not clear, however, that Descartes, who thought that our cause is a non-deceiving God, would have found such a view acceptable. As we saw in Section 3, Descartes’ argument for the truth of his clear and distinct perceptions relies on the premise that when it is psychologically impossible not to judge that P, it is correct to judge that P. As we saw there, Descartes seems to mean “correct” as connoting a positive normative evaluation of a judgment.

Given these three problems, it seems preferable to interpret Descartes as allowing that the atheist has normative certainty in his theorem while he clearly and distinctly perceives it, and that he loses normative certainty only later, when he no longer is attending to the theorem. In order to be normatively certain at this later time, the mathematician would need to know not only that he had a clear and distinct perception at the earlier time, but also that whatever he clearly and distinctly perceives is true. A clear and distinct perception justifies belief. Knowledge that one once had a clear and distinct perception, in contrast, needs to be supplemented by knowledge that one’s clear and distinct perceptions are true.

Descartes’ aim in the Meditations is not only to know some things with certainty at one instant, and other things at another instant. Rather, he hopes to establish the foundations of a science that is stable and likely to last. This requires not only momentary certainty—even justified certainty—but furthermore certainty that can remain intact from one moment to the next. And this requirement for a stable body of knowledge on which to build a lasting science is exactly what the atheist mathematician lacks.

7. Conclusion

I have attributed to Descartes a view of knowledge and of justification according to which you know and are justified in believing whatever you clearly and distinctly perceive, regardless of whether you know the general truth that everything you clearly and distinctly perceive is true. This account of knowledge and justification has something in common with contemporary externalist views, because it is the objective truth-conduciveness of the truth rule, rather than one’s subjective confidence in its truth-conduciveness, which matters for knowledge. But the account also has something in common with contemporary internalist views as well. Because knowledge requires praiseworthiness, and because praiseworthiness requires control, it is central to Descartes’ picture that the only possible candidates for knowledge-conducive rules are...
mental rules. It is thanks to the external guarantee of a non-deceptive God that we are capable of assenting to the truth merely by using our internal faculties correctly. But it is up to us to do so, and it is because of this control that our assent to the truth can be justified and amount to knowledge.
Works Cited


