

Ljiljana Radenovic
Department of Philosophy
Faculty of Philosophy,
University of Belgrade
ljradeno@f.bg.ac.rs
Rastko Jevtic
Institute for Philosophy
Faculty of Philosophy, University of Belgrade
rastko.jevtic@gmail.com

Original Scientific Paper
UDC 14 Декарт Р.
27-167.6

Received: October 14, 2024

Revised: November 20, 2024

Accepted: November 30, 2024



DESCARTES' DEMON AND HIS POWERS: THE BREAK WITH MEDIEVAL DEMONOLOGY

Abstract

In contemporary scholarship, Descartes' Demon has been often treated as a thought experiment with no connection to the world and the views of the 17th century. This rather anachronistic take has been challenged, with some attempting to interpret it within a broader historical Christian context (e.g., Scarre 1990). The most recent challenge has come from scholars focusing on Teresa de Ávila's work (Mercer 2016; Forsman 2023), notably her *Inner Castle*, as it likely had a profound influence on Descartes. Our goal is to show how radically new Descartes' idea of a demon was despite its historical continuity and the apparent influences of monastic literature. We first outline the ontological status of demons and their ways of deception from late Antiquity to Teresa de Ávila and Descartes. Then, we turn to Descartes' Demon and the differences between Descartes' account and traditional ones. While traditional accounts point towards the moral domain (e.g., the ways demons influence humans and the ways we can negate the influence), Descartes' Demon has the purpose of endangering all knowledge. Therefore, Descartes, unlike Teresa de Ávila and other traditional authors, needs a deceiver far more powerful than any demon in the Christian tradition. The attribution of omnipotence to a demon, if only for theoretical purposes, represents a huge break with traditional Christian demonology. Finally, by comparing the Descartes' Demon with the brain-in-a-vat skeptical scenario, we establish two conclusions: (1) the possibility of such a demon represents the most radical form of skepticism, whereas the brain-in-a-vat scenario is less extreme; and (2) these thought experiments demonstrate how underlying metaphysical and theological frameworks shape the formulation of these scenarios and how far they can reach.

Keywords: Descartes · Evil demon · Christian demonology · knowledge · deception

Introduction

A standard practice in introductory courses and textbooks is to treat Descartes's Demon as a decontextualized, "epistemological" thought experiment unrelated to the world and the views of the 17th century. This rather anachronistic take on Descartes has been challenged by scholars, with some attempting to interpret it within the broader historical Christian context (e.g., Scarre 1990). The most recent challenge has been issued by scholars focusing on Teresa de Ávila's work (Mercer 2017; Forsman 2023), especially her *Inner Castle*, as it likely had a profound influence on Descartes. Our main goal is to show that despite its historical continuity and the apparent influences of monastic literature, Descartes' Demon was radically new.

We posit two main comparative theses pertaining to Descartes' Demon and to traditional Christian demonology as a relatively coherent set of doctrines.¹ These two theses indicate the main points of departure between traditional (late Antiquity and Medieval periods) accounts of demons and Descartes' Demon.

1. Powers of traditional demons are limited, but the power of Descartes' Demon is unlimited.

Traditional demons have an ontological place within a hierarchy of beings, with God – a free, omnipotent, omniscient, and benevolent creator – at the top. They are spiritual beings, akin to angels, but their lack of goodness makes them less powerful than their counterparts. Descartes' Demon could not have the same place within the ontological hierarchy. His properties are almost identical to the properties of God; therefore, if he existed, he would take up the top place in the hierarchy, replacing God.

2. Traditional demons are primarily morally significant, but the significance of Descartes' Demon is theoretical, i.e., lacking practical consequences.

Traditional demons create genuine moral obstacles that one needs to overcome daily in order to live a happy and virtuous life. Descartes' Demon serves an exclusively epistemological and metaphysical purpose and has no immediate connection to everyday, practical problems.

The structure of the paper is the following. We begin by outlining the ontological status of demons and their ways of deception from late Antiquity to Teresa de Ávila. We then turn to Descartes' Demon. We argue the difference between traditional Christian accounts of demons and Descartes' Demon reflects the difference between the main worries of the authors: while

1 Even though authors differ in the details of their accounts of demons, we believe the crux of traditional demonological theories is that demons are spiritual beings whose general purpose is to morally violate humans in a rich variety of ways.

Desert and Church fathers, as well as Teresa, focus on the moral domain (e.g., how not to succumb to evil thoughts – the very real threat to every monk and nun), Descartes aims to establish fundamental metaphysical truths, such as the nature of God, thinking, extended substance, etc. For this purpose, Descartes, unlike Teresa and those preceding her, needs a malevolent entity that is incommensurably more powerful than any demon in traditional demonology. This very bold assumption, if only for theoretical purposes, represents a huge break with tradition. Descartes rejects the possibility of such an entity by proving God's existence in the Third Meditation; nevertheless, the existence of the Evil Demon is a real possibility throughout the First and Second Meditations. Our discussion will show how radically new and bold Descartes' idea of the Demon was, despite its historical continuity and the apparent influences of monastic literature. In the end, we aim to achieve two results relevant to contemporary discussions by introducing the brain-in-a-vat thought experiment: (1) to demonstrate that Descartes' Demon is not equivalent to the brain-in-a-vat scenario in terms of the radicality of skepticism – indeed, Descartes' Demon represents a deeper and more radical possibility; and (2) to highlight, through this comparison, the significance of metaphysical and religious background assumptions in shaping the formulation of thought experiments.

1. Short History of Demons: Their Ways of Deception from Early Church and Desert Fathers to Teresa de Ávila

Demons as we know them, i.e., as evil beings with special powers whose sole existence is directed at luring us into eternal damnation, are found in both the Old and the New Testaments in a number of passages (John 8:44; cf. Matthew 4:1-11, cf. Revelation 12:9; 2 Corinthians 4:3-4; 2 Corinthians 2:6-11; Deuteronomy 32:16-17). Satan in the shape of the snake tricks Eve into eating the fruit from the forbidden tree. Satan tempts Jesus in the desert. A number of demons possess both men and women, while Jesus and later the Apostles chase them away (Matthew 8:28-33; cf. Mark 5:1-10; Luke 8:26-39; Matthew 9:32; 12:22; Matthew 12:22; Matthew 4:24; Mark 9:17).

It is interesting to note that early Christians in Late Antiquity who shared their world with pagans believed pagan lesser gods (Daimons) were in fact demons Jesus and Apostles faced and were responsible for all kinds of temptations (Martin 2004). Now, for pagans, Daimons could be both good and evil. Humans needed to make annual sacrifices in temples if they wanted them on their side. For early Christians, all such sacrifices were 'bad faith', i.e., the wrong way of worshiping the Divine, leading them to believe all Daimons were (evil) demons whose goal was to take humans away from God.

Some pagan philosophers, such as Celsus, objected to the widespread belief that Daimons could harm humans. Celsus' argument was that the

moral and ontological orders of the world coincided. Because Daimons were ontologically higher than humans, they could not be morally lower than humans; hence, they could not nor ever would do us harm. This was not the case for Celsus' Christian opponent, Origen. For him, the ontological and moral orders were not the same, and humans could morally supersede otherwise more powerful demons if they had strong faith in Christ. In any case, for early Christians, all Daimons turned into demons, and all demons were simply fallen angels (for more details about the debate between Celsus and Origen, see Martin 2004).

In *City of God*, St Augustine argues pagan gods are literally demons who walk among us with the sole purpose of tempting us to do evil. As evidence for this bold claim, he cites all the morally deprived things these 'gods' are ready to do according to pagans themselves. According to St Augustine, even though demons are fallen angels, i.e., angels who were created by God but rebelled against him, they retain some of their 'superpowers'. This means demons have an advantage over humans to a certain extent. They are faster, can see and hear better, and are able to extract from nature the hidden powers that we humans can only after prolonged research and the slow building of technology. When they make nature work 'faster' (e.g., when they make flowers bloom rapidly), this may appear to be a miracle. But according to Augustine, this is not a true miracle, as all the seeds of what can happen in nature were planted by God inside the things. The true miracle is God's creation; the rest is appearance. Demons can deceive us with these appearances only because we are slower than they are (see Cameron 2010).

Demons played a very important role in monastic literature as well. In the lives of the Saints, we see how often such profoundly Christian figures were exposed to demonic temptations. Their sanctity was understood to be demonstrated in their ability to discern and fight demons and to chase them away just like Jesus and the Apostles did. Stories of their powers against demons are found in many places, including *Life of Antony* by Athanasius, *Lausiac History* by Palladius, and *Conferences* by Cassian. Evagrius' *Eight Evil Thoughts* gives details on the kind of thoughts and passions demons could invoke in monks (and other humans as well) and the ways to fight their temptations.

Let us briefly explore the kind of thoughts demons could invoke and the mechanisms they used.² According to Evagrius, demons wait patiently for the monk's moment of weakness, and as soon as they recognize it, they plant one or several evil thoughts in him. Once these thoughts are planted, they multiply, invoke other equally bad thoughts, and eventually lead the monk astray from his godly path. For instance, overeating (the sin of gluttony) yields lustful imaginings (the sin of fornication) and from here, other evil thoughts can arise: avarice, anger, sadness, and the like. If the monk is exhausted from fasting, the demon can convince him that it is not that bad to overeat after

2 For more details see Sorabji (2000, pp. 343-372).

all. Strictly speaking, the monk is never safe. Even if he successfully fights the demon and never succumbs to evil thoughts about gluttony, the demon of pride and vainglory is there to make him think he is better than anybody else. It is important to note that demons have the power to plant evil thoughts in the monk and to deceive him into believing he is not doing anything wrong – that the gluttony or fornication he is about to commit is not that bad after all. The monk has no control over this initial temptation, but it is in his power to resist it and not indulge in it. In other words, the monk is free to sin or not to sin despite the demon's 'indecent proposals'.

Furthermore, when the monk is besieged by demons who whisper ungodly thoughts, he can try to find the remedy for his situation. However, it is in the demon's power to suggest the wrong remedies. So if the monk feels sad (the demon has already won so to speak), he might start looking for ways to feel better. When he does so, the demon is ready to suggest other ungodly things to feel better (to go back to his previous life, to chitchat and gossip with fellow monks, to leave his cell and move to another place, etc.). If he falls for the wrong remedy, he falls further into the spiral of sin. Thus, it is extremely important for the monk to learn to recognize false and true remedies. Evagrius argues if the remedy is basically self-serving and ego pleasing, it is certainly the devil's work. But if the remedy encourages self-sacrificing and love of God, then it is the true one (for more details, see Radenovic & Akkad 2022).

To sum up, in Late Antiquity, demons were real; they roamed the earth and meant us no good. Closer to Descartes' time, into the high Middle Ages and the time of Thomas Aquinas, demons were still around and their main job was to daunt humans. According to Aquinas, both angels and demons are spiritual beings, and just as Augustine argued, they were created by God. There is a hierarchy among them, with some angels given more grace and power than others. For demons, the hierarchy remains even in their Fall, so some remain more powerful. Both angels and demons can use the natural potential of objects to manipulate nature and cause surprising (to us!) things to happen. Indeed, demons use these powers for their own evil goals and to deceive humans. Neither angels nor demons can bend human will directly. They can manipulate external natural forces and make something that looks like a miracle happen in front of human eyes. Alternatively, they can present to the human mind images via senses (or by moving internal humours in a particular way) to nudge humans to do a specific thing. For Aquinas, demons are not capable of 'planting' evil thoughts directly into the human mind, but they can definitely lead us to them; i.e., they can confuse us and lead us to conclude that what is bad is good and vice versa (Cameron 2010).

With this brief outline in mind, and having seen the powers demons were given throughout the long period between early days of Christianity and the late Middle Ages, let us briefly turn to Early Modern period and

Teresa de Ávila's demons. In 2017, Christia Mercer published a seminal paper on the highly probable influence of Teresa de Ávila's work *The Inner Castle* on Descartes' *Meditations*. In 2023, Jan Forsman further developed Mercer's ideas by highlighting similarities between Descartes' Demon and Teresa de Ávila's evil demons and the ways of their deception. Mercer and Forsman agree there is no direct evidence that Descartes read Teresa de Ávila's work, but they suggest he probably did since she was a 'best seller' of the time and a popular read among Jesuits. They convincingly show the structure of Descartes' *Meditations*, as well as the steps of increasing doubt, follow Teresa's manuscript. Still, while the parallels between these two works are remarkable and undeservedly forgotten, Teresa de Ávila's goal, just like the goals of the Desert fathers of the 4th century, was to help her fellow nuns to avoid temptation and remain firmly on their godly path. Like Evagrius, Teresa wanted to delineate the thoughts that were about true goodness, moral behaviour, godly feelings, and those that only appear to us as such but are the devil's temptations. Obviously, the devil could confuse monks and nuns about what was good and bad, and they needed to find the way out.

So how did demons deceive monks and nuns according to Teresa? She says that as soon as we think we know the difference between good and evil, demons make us uncertain about what we know and lead us back into confusion. For instance, as soon as the nun (in this case, Teresa herself) is convinced it is good to leave the world behind, along with all the people, wealth, etc., demons revive her love for the world and "bring to mind the esteem one has in the world, one's friends and relatives, one's health... the devils represent these snakes (worldly things) and the temporal pleasure of the present as though almost eternal" (II: 1, 3 (299) as quoted in Mercer 2017). As we can see, Teresa's view on demons remains within Christian tradition, in that demons have limited powers and are eager to confuse nuns about moral matters so they keep on sinning. This is neither the nature nor the purpose of Descartes' Demon, as we discuss next.

2. Descartes' Demon

It is commonly understood that the job of Descartes' Demon is to put the Meditator into the worst epistemological situation imaginable. However, why Descartes uses a demonic being to accomplish this goal and how it happens are not obvious. To clarify these choices, it is necessary to carefully follow the development of the First Meditation – the most important source of information on Descartes' Demon (AT VII 20, 21; CSM I 14). By doing so, we hope to support the thesis that Descartes' Demon is unlimited and has exclusively theoretical relevance.

The idea of the Demon is introduced in the First Meditation after it is established that the knowledge of the simplest and most general things resists

the skeptical possibility of being a dream. In other words, arithmetic and geometry (i.e., "corporeal nature in general, and its extension; the shape of extended things; the quantity, or size and number of these things; the place in which they may exist, the time through which they may endure, and so on" (AT VII 20, 21; CSM II 14)) are not made dubitable on the basis of the possibility of the Meditator dreaming. The stage is set for the final phase of doubt – an even worse possibility is needed to make all the mentioned things uncertain. The following passage opens the door to such a possibility and marks the first step towards the introduction of Descartes' Demon; we will call it the *first passage*:

And yet firmly rooted in my mind is the long-standing opinion that there is *an omnipotent God* who made me the kind of creature that I am. How do I know that he has not brought it about that there is no earth, no sky, *no extended thing*, no shape, no size, no place, while at the same time ensuring that all these things appear to me to exist just as they do now? What is more, since I sometimes believe that others go astray in cases where they think they have *the most perfect knowledge*, *may I not similarly go wrong every time I add two and three or count the sides of a square, or in some even simpler matter, if that is imaginable?* But perhaps God would not have allowed me to be deceived in this way, since he is said to be supremely good. But if it were inconsistent with his goodness to have created me such that I am deceived all the time, it would seem equally foreign to his goodness to allow me to be deceived even occasionally; yet this last assertion cannot be made. (AT VII 20, 21; CSM I 14; our italics).

The deception that an omnipotent, supreme deceiver could accomplish is far beyond the deception we would find ourselves in if we were dreaming. Descartes raises the idea of such a being to radicalize doubt and make the simplest and most general things questionable.

Now, let us look at the next passage where Descartes writes about the supreme deceiver for the second time; here, the Demon is introduced in its completeness. We will call this the *second passage*:

I will suppose therefore that *not God*, who is supremely good and the source of truth, but rather some *malicious demon* of the utmost power and cunning has employed all his energies in order to deceive me. I shall think that the sky, the air, the earth, colours, shapes, sounds and all external things are merely the delusions of dreams which he has devised to ensnare my judgement. I shall consider myself as not having hands or eyes, or flesh, or blood or senses, but as falsely believing that I have all these things. I shall stubbornly and firmly persist in this meditation; and, even if it is not in my power to know any truth, I shall at least do what is in my power, that is, resolutely guard against

assenting to any falsehoods, so that the deceiver, however powerful and cunning he may be, will be unable to impose on me in the slightest degree. (AT VII 22, 23; CSM I 15; our italics)

These two passages are the most important passages in the whole of Descartes' corpus when it comes to the worst possible epistemological situation: the existence of a supreme being who is also a deceiver. When we follow the text carefully, we understand that the idea of the Demon does not enter the scenario by itself, independently of ideas of other possible entities which, *prima facie*, could do the same. The idea of a *malicious demon* found in the *second passage* enters the text as a replacement for the idea which comes chronologically first: the idea of God. Thus, at least narratively, the notion of Descartes' Demon is derived from or dependent on the idea of God.

Narrative dependence is a facade for a deeper kind of dependence. Let us investigate the notion of dependence of Descartes' Demon on God and answer the following question: on what basis does the replacement happen at all? The replacement of one entity by another cannot happen if the two have nothing in common; they must have a key property in common (or even more than one) for the replacement to happen. This is why we need to unpack the logical relationship between the idea of a malicious Demon and God. We believe the most important analogical and disanalogical properties can be found in the two passages cited above. The disanalogical property is the main reason for the replacement of the idea of God by the idea of the Demon.

Let us start with the analogical property. In the first sentence of the *first passage*, Descartes explicitly mentions omnipotence as a relevant property of God: "firmly rooted in my mind is the long-standing opinion that there is *an omnipotent God* who made me the kind of creature that I am". Following this, in the *second passage* in which the replacement happens, Descartes uses a synonymous expression we italicize for emphasis: "I will suppose therefore... some malicious demon *of the utmost power*." Descartes needs to establish a doubt about the possibility that is more powerful than the possibility of being in a dream. What other possibility could be worse than the deception of the omnipotent creator? Why would Descartes explicitly attribute omnipotence to God and to the Demon, if it wasn't the key analogical property?

Our claims that the notion of God is the notion from which the notion of the Demon is derived and that omnipotence is a key analogical property are strongly supported by the second most important formulation of the most radical phase of doubt in all of Descartes' texts. In the 5th paragraph of Principles of Philosophy I, Descartes writes:

One reason for this (doubt in demonstrations of mathematic and self-evident principles; our clarification) is that we have sometimes seen people make mistakes in such matters and accept as most certain and self-evident things which seemed false to us. Secondly, and *most*

importantly, we have been told that there is *an omnipotent God who created us*³. Now we do not know whether he may have wished to make us beings of the sort who are always deceived even in those matters which seem to us supremely evident. (AT VIII A 6; CSM I 194; our italics)

Here, Descartes does not replace God with a malicious demon explicitly, but he talks about an omnipotent God who could create us as beings who are deceived in all matters. Deception is definitely contradictory to the benevolent nature of God, and Descartes is certainly going to argue that God is necessarily good. The crucial thing to understand about the 5th paragraph is that Descartes does not treat the notion of God as fully developed at this moment; thus, the contradiction is not acknowledged, and the possibility of the deceitful supreme creator is 'superficially' allowed. Nonetheless, the interweaving development of the notions of God and the Demon in the same context is enough to establish the analogy between them.

What is omnipotence? What does it consist of? Such a property cannot be easy to define and we do not wish to clog up the paper with a discussion on metaphysical issues – this is simply not our purpose. Therefore, we suggest a definition we believe is illuminating, even though we don't plan on defending it from extra-textual objections. Omnipotence is a property that consists of (a) the possibility to create freely and (b) the possibility to create *ex nihilo*. The Christian God is omnipotent by this definition, in contrast to pagan gods and in contrast to traditional demons.

Descartes' emphasis on creation as a property shared by God and the Demon supports our definition of omnipotence as the key analogical property. In the *first passage*, Descartes speaks about an entity which has made him the way he is and which can bring about other things. The emphasis on creation is obvious in Principles I, as well: in the 5th paragraph (“...most importantly, we have been told that there is *an omnipotent God who created us*.”) and the 6th paragraph (“But whoever turns out to have created us...”).

The *second passage* we quoted from Meditations is a really interesting case. An inattentive reader could easily miss the emphasis on creation and read the passage as if the emphasis is on deception *simpliciter*. However, deception itself is a kind of creation. Descartes' Demon deceives by creating ideas which have no correspondent entities in reality. In other words, the difference between creation and deception is that deception is the creation of false ideas in the mind, while creation is a general term for the actualization of any possible thing.

In conclusion, a property of omnipotence which consists of the possibility of free creation out of nothing, i.e. without any limits, is shared by God and Descartes' Demon. As we have argued in the previous section, this property does not traditionally belong to demons.

3 Our definition of omnipotence and the claim that it is the main analogical property is warranted by Descartes' emphasis on creation in this passage as well.

Now we move to the disanalogical property. As soon as the replacement of God by the Demon is understood to be happening, the disanalogical property becomes obvious. To remind readers: Descartes prepares the introduction of the Demon by writing: “perhaps God would not have allowed me to be deceived in this way, since he is said to be supremely good”. In the *second passage* he replaces God with the Demon: “I will suppose therefore that not God, who is supremely good and the source of truth, but rather some *malicious* demon of the utmost power and *cunning* has employed all his energies in order to deceive me.”

It is crystal clear that Descartes is replacing one omnipotent entity with another; the twist is that the replacement entity is immoral. While God must be benevolent because God is the (source of) good itself and as such cannot create false ideas without (good) reason, the replacement entity must be malicious because Descartes needs to evoke the worst kind of doubt possible. And where should a good Christian look for an evil spirit to replace a spirit that is supremely good? From a commonsensical Christian perspective, the answer is as clear as a day. Demons are understood as paradigmatic evil spirits throughout Christian history, and Descartes’ choice of a replacement is straightforward and intuitive.

Descartes’ own formulation contains both key analogical and key disanalogical properties, thus summing up our view: “[But] there is a deceiver of supreme power and cunning who is deliberately and constantly deceiving me.” The deceiver is both supremely powerful and cunning; at this moment in the text, this is synonymous with the claim that the deceiver is omnipotent and malevolent.

Finally, the most explicit and telling formulation of the replacement of the idea of God by the idea of the Demon briefly appears at the start of the Third Meditation (AT VII 36; CSM II 25): „Since I have no cause to think that there is a *deceiving God*, and I do not yet even know for sure whether there is a God at all, any reason for doubt which depends simply on this supposition is a very slight and, so to speak, metaphysical one” (our italics). This passage makes it obvious that Descartes is modeling the notion of the Demon on the notion of God. Since the Demon is introduced into the text by being juxtaposed with God, we can assume the Demon inherits all the properties of God except benevolence. The most important properties in our context are omnipotence, omniscience, and absolute freedom. Traditional demons definitely lack these qualities.

One could think of the incomplete notion of God as a sketch that Descartes slightly modifies from one passage to another, in relation to the context of the passage. At first, he varies certain properties of such a being, but the final result is a total change of identity and a new being. In the *first passage*⁴ (quoted from the First Meditation), Descartes suspends the analytical consequences of the

4 See p. 145.

notion of benevolent God to fully express the epistemological consequences of the lack of such a being. The same is true for the 5th paragraph of Principles I⁵: Descartes pretends not to see that deception results from the lack of benevolence, and as such, is contradictory to God. In the 6th paragraph of Principles I, Descartes fully recognizes this contradiction, ceasing to call such a being God ("But *whoever* turns out to have created us, and however powerful and however deceitful he may be..."; our italics). However, the *second passage*⁶ goes a step further: the contradiction is not only recognized and accepted, but the property of benevolence is understood as essential to God. If the property is essential, it cannot be varied/removed without changing the identity of the being in question; therefore, the notion in question stops being the notion of God and becomes the notion of a malicious demon. Throughout all these changes, the property of omnipotence is attributed as the most important analogical property of the relevant beings, along with the capacity for free creation, out of which omnipotence emerges.

To better understand omnipotence as a property that differentiates Descartes' Demon from traditional demons, we should answer the following question: on what objects does the Demon exercise his powers? Descartes' list of the things we may be deceived about is long and unsystematic. He mentions earth, sky, shape, size, place, colours, and sounds, as well as our own hands, as eyes, flesh, blood, and senses (AT VII 22, 23; CSM I 15). Even though the list is inexhaustive⁷, expressions such as "all external things" and "senses," as well as the general philosophical context and the purpose of radical doubt, lead us to the only plausible and non-controversial conclusion: the objects of deception are *all of the ideas of external things*.⁸

The gap between Descartes' Demon and traditional demons becomes even bigger when we remind ourselves that the objects of Descartes' Demon include mathematical truths. His omnipotence is manifested even in the cases of the most perfect knowledge such as mathematics ("*May I not similarly go wrong every time I add two and three or count the sides of a square, or in some even simpler matter, if that is imaginable?*"; AT VII 20,21; CSM I 14; our italics). In other words, the supreme deceiver may also deceive us when it comes to the simplest mathematical operations, such as arithmetical and geometrical addition. We don't find traditional demons with the power to do this. Therefore, all the ideas of external things and all the ideas of mathematics are victims to the vile doings of the omnipotent Demon; traditional demons are not capable of such deception.

5 See pp. 146–7.

6 See pp. 145–6.

7 Descartes is not to be blamed for the lack of systematicity in these parts of the text. After all, he has just started meditating, and his categorical apparatus is not developed. For example, sense perception is the only faculty on the list that is not an object or property.

8 This is precisely why Descartes makes an effort to prove the existence of external world.

Therefore, Descartes' Demon has the purpose to deceive the Meditator about all of his ideas.⁹ In retrospect, after we are done reading the Meditations, we know this is not possible – certain ideas are clear and distinct, and we cannot be mistaken about them. However, at the end of the First Meditation – at the height of the Demon's power – universal deception is a real possibility. This stands in stark contrast to traditional demons that always deceive about a particular, this or that idea. In other words, the malicious Demon's influence is universal and infinite, not particular and finite. The situation differs greatly from the traditional situation: even if traditional demons could make an idea of something represent nothing, they couldn't do so with every idea.

Mirroring Descartes' strategy, we have juxtaposed God and the Demon to argue for the unlimited nature of Descartes' Demon and its sole theoretical relevance. Now, we want to add a few more remarks to solidify our contention that Descartes' Demon has a metaphysical and epistemological purpose, not a moral and practical one.

First, and most importantly, there are no textual indications that Descartes' Demon serves any philosophical purpose outside "First philosophy". In other words, no mention of the Demon's influence can be found in any other part of Descartes' system except his metaphysics and epistemology. Strictly speaking, the Demon's possible influence is eliminated as soon as the existence of God is proven, and the effects of this proof keep surfacing until the end of the Meditations. Descartes' Demon simply stops being relevant when one has finished meditating.

Descartes describes the doubt that culminates in the possibility of the Demon existing by using the words "metaphysical" and "hyperbolic"¹⁰ and goes on to explain this characterization. In a passage at the beginning of the Third Meditation (AT VII 36; CSM II 25), he writes: „Since I have no cause to think that there is a deceiving God, and I do not yet even know for sure whether there is a God at all, any reason for doubt which depends simply on this supposition is a very slight and, so to speak, *metaphysical one*" (our italics). Since the most extreme phase of doubt consists of the possibility of the evil Demon existing, it is intuitive to characterize the Demon as "hyperbolic". However, in light of our research, we argue there is another, deeper and more interesting reason to call Descartes' Demon hyperbolic: it is modeled on God, not on any kind of traditional demon. This feature makes the Demon hypothetical as well – if God didn't exist, then the Demon could actually exist – while traditional demons were very real for a genuine, believing Christian.

9 Sometimes, the truths of logic are also mentioned as the targets of the Demon. Unfortunately, we don't have enough space to elaborate on the complexity of this issue.

10 CSM translation uses the word "exaggerated", which is in a way an even more informative choice than Descartes' choice.

For the sake of Descartes' Modernity, it is worth mentioning that traditionally conceived demons are nowhere to be found doing explanatory work in the whole of his system. The only kind of demon left is the (hyperbolic) Demon. Even in the domain where we should be expecting the workings of traditional demons – the domain of psychology and ethics – Descartes approaches the topic naturalistically, as a scientist. The whole of *Passions of the Soul* bears witness to this: no psychological entity or process is explained by the causal influences of traditional demons. Bad thoughts and actions are explained naturally; demons become redundant or worthless as explanatory entities in Descartes' practical philosophy. Descartes' explanation of *akrasia* is a great example of this – there is no sign of any influence of spiritual beings that results in the “weakness of the will.”¹¹

Finally, we want to explicate the single most important consequence of our analysis for contemporary epistemological debates on skepticism. We argue that Descartes' modeling of the Demon on God results in the most radical possible skeptical scenario and that there can be no other scenario imaginable which is as radical. We will compare Descartes' Demon to the popular brain-in-a-vat thought experiment¹² in order to show that there is a crucial ontological difference between them. The ontological difference gives rise to a difference in the skeptical reach of these thought experiments.

Let us first briefly describe the brain-in-a-vat scenario in order to explain the ontological difference. Of course, one must turn to Putnam's setup of the thought experiment since he is the one who made it widely popular:

“Here is a science fiction possibility discussed by philosophers: imagine that a human being (you can imagine this to be yourself) has been subjected to an operation by an evil scientist. The person's brain (your brain) has been removed from the body and placed in a vat of nutrients which keeps the brain alive. The nerve endings have been connected to a super-scientific computer which causes the person whose brain it is to have the illusion that everything is perfectly normal. There seem to be people, objects, the sky, etc; but really all the person (you) is experiencing is the result of electronic impulses travelling from the computer to the nerve endings. The computer is so clever that if the person tries to raise his hand, the feedback from the computer will

11 For Descartes' explanation of *akrasia*, see paragraph 47 of the first part of *Passions of the Soul*.

12 Recently, there has been considerable discussion about the possibility that we are living in a simulation. We believe that for all relevant epistemological purposes, the brain-in-a-vat thought experiment can be understood as equivalent to the simulation hypothesis. It is impossible to discuss the simulation scenario without referencing the Matrix which is undoubtedly the paradigmatic example of this idea. Our thesis is supported by Chalmers' claim that “in effect, Neo is a brain in a vat” (see “The Matrix as Metaphysics”, p. 132, *Philosophers Explore the Matrix*, ed. by C. Grau) We will simply write about the brain-in-a-vat in this paper since it predates *The Matrix* and is discussed in well-known literature.

cause him to 'see' and 'feel' the hand being raised. Moreover, by varying the program, the evil scientist can cause the victim to 'experience' (or hallucinate) any situation or environment the evil scientist wishes. He can also obliterate the memory of the brain operation, so that the victim will seem to himself to have always been in this environment. It can even seem to the victim that he is sitting and reading these very words about the amusing but quite absurd supposition that there is an evil scientist who removes people's brains from their bodies and places them in a vat of nutrients which keep the brains alive. The nerve endings are supposed to be connected to a super-scientific computer which causes the person whose brain it is to have the illusion that. . ." (Putnam 1981 p.6)

In the following passage, Putnam mentions a way to interpret this thought experiment: "When this sort of possibility is mentioned in a lecture on the Theory of Knowledge, the purpose, of course, is to raise the classical problem of scepticism with respect to the external world in a modern way. (How do you know you aren't in this predicament?)" (ibid.) If we accept Putnam's claim that the brain-in-a-vat scenario invokes skepticism about the knowledge of the external world, then the main difference between this scenario and Descartes' Demon becomes fairly obvious. As we already mentioned, an important consequence of the possibility of the Demon's influence is the universality of doubt: it's not only that the knowledge of external world is put into question, but the knowledge of arithmetic and geometry as well. Descartes' Demon has the purpose to instill doubt not only about the ideas of the external world, but about every idea meditator can have.

Be that as it may, we believe that there is an even more important difference which stays hidden at a first glance. The evil scientist¹³ which is by assumption, in charge of brain-in-a-vat system is nonetheless a finite being! This being may be extremely powerful, but it is far from being omnipotent! The limitation of power makes for a crucial difference between the brain-in-a-vat scenario and Descartes' Demon. Remember our definition of unlimited power: "Omnipotence is a property that consists of (a) the possibility to create freely and (b) the possibility to create *ex nihilo*."¹⁴ The component (b) is the one we focus on to evade irrelevant complications about free will. However powerful the evil scientist, he cannot create the experiences the brain-in-a-vat has *ex nihilo*. In other words, he cannot create the ideas which do not represent anything – have no correspondents in reality – out of nothing. The only option left is that there must be something which grounds the experiences the brain-in-a-vat has – some kind of basis for the simulation. What this means means is that there is something real about these experiences after all!

13 Sometimes the evil scientist is replaced by aliens or machines. This does not matter: the point about finiteness of the deceiving entity will necessarily stand in order for the thought experiment to retain identity.

14 See p. 147.

It will be very useful to rewind to the skeptical possibility of dreaming for a moment. Remember that Descartes introduced the Demon in order to destroy knowledge that is left over after the possibility of being in a dream. We propose to read following passages while thinking about the possible similarities between being a brain-in-a-vat and dreaming. We will italicize the key moments:

"Suppose then that I am dreaming, and that these particulars — that my eyes are open, that I am moving my head and stretching out my hands — are not true. *Perhaps, indeed, I do not even have such hands or such a body at all.* Nonetheless, *it must surely be admitted that the visions which come in sleep are like paintings, which must have been fashioned in the likeness of things that are real, and hence that at least these general kinds of things — eyes, head, hands and the body as a whole — are things which are not imaginary but are real and exist.* For even when painters try to create sirens and satyrs with the most extraordinary bodies, they cannot give them natures which are new in all respects; they simply jumble up the limbs of different animals. Or if perhaps they manage to think up something so new that nothing remotely similar has ever been seen before — something which is therefore completely fictitious and unreal — at least the colours used in the composition must be real. By similar reasoning, although these general kinds of things — eyes, head, hands and so on — could be imaginary, *it must at least be admitted that certain other even simpler and more universal things are real.* These are as it were the real colours from which we form all the images of things, whether true or false, that occur in our thought.

...

So a reasonable conclusion from this might be that physics, astronomy, medicine, and all other disciplines which depend on the study of composite things, are doubtful; while *arithmetic, geometry and other subjects of this kind, which deal only with the simplest and most general things, regardless of whether they really exist in nature or not, contain something certain and indubitable. For whether I am awake or asleep, two and three added together are five, and a square has no more than four sides*¹⁵. It seems impossible that such transparent truths should incur any suspicion of being false." (AT VII 19; CSM II 13; our italic)

The essence of the dream scenario is the claim that all concrete experiences of external things may be false representations – since there is a possibility that the concrete objects they represent do not exist (even the experience of

15 On the basis of these claims it is obvious that Descartes' wouldn't allow the dreaming scenario to disrupt the mathematical truths. We have provisionally allowed for this to happen in the case of the brain-in-a-vat scenario (although we don't find it intelligible) in order to accentuate the difference between the infinite and finite deceiver.

our body may be false). This is precisely the essence of the brain-in-a-vat scenario as well. Therefore, the two scenarios are identical when it comes to their epistemological status. For example, both “a sleeper” and “a brain” may have experiences of “its” body while no body actually exists. More generally, both “a sleeper” and “a brain” may have experience of any concrete thing while this experience does not represent anything.

However, Descartes believes that experiences in these situations cannot be altogether representationally worthless. Even though these experiences, *in concreto*, do not represent anything, they must contain simple elements that represent something existent. It is these elements that constitute the knowledge which is not affected by the skeptical hypothesis of dreaming or being a brain-in-a-vat. Most simple and general truths are not affected by the dream scenario, since dreams are like paintings: they must contain true simple elements – elements that represent something real – even though these elements may not represent reality when they are combined into a painting as a whole.

Descartes’ description of a dream is analogous to the description of the situation that brain-in-a-vat is found in. Therefore, Descartes’ diagnosis of the skeptical reach of the scenario must also be analogous: he continues looking for a more radical possibility that can endanger the most simple and general truths as well. In other words, Descartes needs a more radical option and he finds it in the possibility of the deceiving God. It is exactly after the quoted passages that the idea of the Demon is introduced in order to overcome the “dream stage” of skeptical doubt.

Since we believe the analogy between dreaming and brain-in-a-vat is properly established, we can reformulate Descartes’ assumption about the “leftover” knowledge of the simple and general natures in a more contemporary framework. His assumption can be “modernized” by claiming that there must be real basic elements of experience out of which the evil scientist builds the hallucinatory experience/simulation (on the “dualistic” interpretation of the brain-in-a-vat thought experiment) or real material constituents which are the foundation of the hallucinatory brain processes (on the material interpretation of the thought experiment). While we try to evade the extremely complicated problem of mind-body gap with these formulations, it is clear that the skeptical scenario just isn’t radical enough – there is something real that can be known while being in a dream or a brain-in-a-vat. This something cannot be known if a subject is under the influence of the Demon, i.e. a deceiving God.

By making the comparison with the brain-in-a-vat thought experiment we wanted to achieve two things. First, we wanted to show that because Descartes’ Demon inherits all of God’s properties with the exception of benevolence, the possibility of his existence is the most radical skeptical scenario possible. The worst possible epistemological situation one can be in is the situation in which the infinite and omnipotent creator has the purpose

to deceive a finite being. Such a scenario is crucially different in comparison to the brain-in-a-vat scenario in which the deceiver is very powerful, although infinitely and incommensurably less so than God. The evil scientist has limitations when it comes to his creative power – he cannot create out of nothing, but his deceptions need to have real causes – and this is a crucial difference which makes for a less radical skeptical possibility. This is also what makes the brain-in-a-vat analogous to the dreaming scenario.

The second point which surfaces out of our comparison is the general dependence of thought experiments on the metaphysical and religious framework. A framework, although frequently indeterminate i.e. lacking both concrete and abstract details¹⁶ contains certain characteristics which influence the way the thought-experiment is staged. In other words, these characteristics are the fertile ground on which the thought experiments flourish. The Cartesian backdrop is early modern, although still Christian; the brain-in-a-vat framework is naturalistic – it describes a possibility that is still in the realm of science-fiction, although one day it could be made possible. The Cartesian framework has an omnipotent (and sometimes malicious) God as its centerpiece; on the other hand, the most natural background for the brain-in-a-vat scenario is the physicalistic metaphysics which allows for a deceiver that is only a part of the world. The Cartesian framework enables for the formulation of a more radical scenario – a scenario that contains a transcendent being; unfortunately (or luckily!) for contemporary epistemology, such an option is hard to imagine inside a naturalistic framework.

Ultimately, Descartes' Demon can be called a radical thought experiment, albeit one embedded in Christian tradition. It is a philosophical venture rooted in Christianity, at the same time delving into unimaginable territories from the perspective of a traditional Christian framework. As such, it is a venture of the 17th century: steeped in the medieval worldview while giving birth to Modernity.

References

- Cameron, E., (2010). *Enchanted Europe: superstition, reason, and religion 1250-1750*. Oxford University Press.
- Chalmers, David J. (2005). "The Matrix as Metaphysics". In Christopher Grau (ed.), *Philosophers Explore the Matrix*. Oxford University Press. pp. 132.
- Descartes, R. *Oeuvres de Descartes*, 11 vols., ed. Adam, C., Tannery (AT), P. Paris: Vrin/CNRS, (1964–76);

16 Aren't these properties exactly the reason it possesses a creative force, stimulating the formulation of thought-experiments which contribute to the exploration of the significance and the intellectual reach of the framework?

- Descartes, R. *Philosophical Writings of Descartes*, 3 vols., trans. Cottingham J., Stoothoff, R., Dugald Murdoch, D. and Kenny, A. Cambridge University Press, (1984-91);
- Forsman, J., (2023). "Teresa's Demons: Teresa of Ávila's Influence on the Cartesian Skeptical Scenario of Demonic Deception". *Journal of the History of Women Philosophers and Scientists*, 1(aop), pp.1-21.
- Martin, D. B. (2004). *Inventing superstition: From the Hippocratics to the Christians*. Harvard University Press.
- Mercer, C., 2017. "Descartes' debt to Teresa of Ávila, or why we should work on women in the history of philosophy". *Philosophical Studies*, 174, pp.2539-2555.
- Putnam, H., (1981). *Reason, Truth and History*. New York: Cambridge University Press.
- Radenovic, L. and Akkad, I., (2022). "History of Emotional Suffering: from emotions to needs in the history of emotions". *History and Theory*, 61(1), pp.96-123.
- Scarre, G., (1990). „Demons, demonologists and Descartes“. *The Heythrop Journal*, 31(1), pp.3-22.
- Sinkewicz, R.E. (2003) ed. and transl., *Evagrius of Pontus: the Greek ascetic corpus*, Oxford Early Christian Studies, New York: OUP
- Sorabji, R., (2000). *Emotion and Peace of Mind, from Stoic Agitation to Christian Temptation*. Oxford University Press.