Demon Skepticism and Non-Veridical Concepts

Claude Panaccio perfectly reconstructed my position and my argument he is criticizing. In fact, I have nothing to complain about his interpretation of my idea of a BIV or of non-veridical concepts, up to a certain point. To see exactly where our interpretations of the relevant issues diverge, it will be useful to take a closer look at the idea of “Demon Skepticism” itself.

In The Matrix, the celebrated movie premised on a brains-in-a-vat scenario, there is an interesting conversation among “the rebels”, i.e., persons living in the devastated physical world of the 22nd century, who originally acquired their concepts in the virtual reality of “the Matrix”, a computer program feeding artificially generated humans, nurtured in complete sensory isolation from physical reality, the virtual experiences of 21st century America as we know it. The conversation concerns what the artificial peptide goo served for dinner tastes like. The suggestions range from runny egg to Tasty Wheat to snot. But it soon turns out that the main concern is not that one of the interlocutors makes an error in judgment in the sense that what he deems, say, Tasty Wheat taste is really oatmeal, or chicken, or tuna taste. Rather, the concern is that the interlocutors literally have no idea of Tasty Wheat taste or chicken taste or tuna taste, or of genuine chicken or real tuna, for that matter. Having acquired their concept of, say, chickens in the virtual reality of the Matrix, in complete cognitive isolation from a real world that at least used to be populated by real chickens, this concept can only represent the virtual objects of this virtual reality, whatever those are, but not the real objects of physical reality.

But if we come to think about it, Descartes had the same sort of concern when he worried not only about the possible formal falsity, as he called it, of our judgments, but also about the material falsity of our simple ideas, whether they are simple sensory or intellectual ideas, and whether these intellectual ideas are adventitious, empirical ideas, acquired somehow from sensory ideas, or innate intellectual ideas, co-created with our minds. For Descartes is not so much concerned about the sort of relatively easily corrigible error in judgment that stems from ordinary sensory illusion (after all, that’s how we know about sensory illusion at all, namely, by our ability to detect it!), as about the systematic, in principle incorrigible error stemming from the material falsity of our simple ideas, whether they are acquired in a scenario of systematically deceptive quasi-experiences envisioned in the “dreaming argument” or planted in our minds by its maker or an omnipotent manipulator, as envisioned in the “demon argument”.

Descartes’ way out of the epistemic predicament of the Demon scenario is to show that it is not really possible: at least one of my simple ideas must be materially true, namely, the idea of myself; for otherwise I would have to accept the obviously self-defeating claims that I do not doubt, I do not think, I do not exist, whereas all these are refuted by the very act of doubting everything on account of thinking about the alleged possibility of the Demon scenario. Having proved the material truth of the idea of ego, Descartes moves on to prove the material truth of the idea of God, by proving God’s existence, eliminating the possibility of there being an omnipotent deceiver, which in turn proves the reliability of clearly and distinctly perceived truths of reason. From this, namely, the formal truth of principles of reason, Descartes infers the material truth of all innate ideas making up these principles, which is the guarantee of their applicability in all possible creations, but which still leaves open the doubt concerning the empirical reality of an actual creation on account of the dreaming scenario. Descartes’ crucial move, then, is to use the material truth of innate ideas and the identity of the formal objects of these ideas in all possible creations and the formal objects of some of our sensory ideas in this creation to infer that the
latter also have to be materially true, adequately representing objects of an empirical, actual physical reality.

This little, somewhat tendentious, recapitulation of the material to be covered in any first year history of philosophy seminar merely served to illustrate two important points with regard to the idea of Demon skepticism as I understand it. First, Demon skepticism concerns not only doubt concerning the formal truth of our judgments, but also doubt concerning the material truth of our simple ideas, or, in other words, the veridicality of our simple concepts. Second, since according to Descartes, the formal truth of judgments entails the material truth of the ideas making them up, by contraposition, he takes it to be self-evident that the material falsity of our ideas entails the formal falsity of the judgments they make up, provided this material falsity is understood as the complete failure of these ideas to engage reality, which is the way I would interpret my description of a non-veridical concept.

In fact, semantically speaking, we may clarify the idea of a non-veridical concept’s failure to engage reality by saying that in a formal semantics, categorematic terms expressing non-veridical concepts and those expressing veridical ones would take their semantic values from two disjoint sets, even if, perhaps, phenomenally, from the perspective of the minds that form these concepts, they may be indistinguishable. Thus, for instance, if I have the concept of chickens formed in physical reality upon encountering genuine chickens, then the concept I express by the word ‘chicken’ represents genuine chickens. On the other hand, if I was raised in the Matrix, what I can express by the word ‘chicken’ is at best a concept that represents virtual chickens, whatever those are, but definitely not chickens as we understand them. Still, the claim of the Demon argument is that I can have phenomenally the exact same mental contents whether I acquire my concepts in genuine or in virtual reality. So, the concept acquired in the Matrix would appear to me to represent the same in the same way as the concept acquired in genuine physical reality, despite the fact that only the latter represents chickens, and the former does not. It is this idea of a non-veridical concept that I briefly described by saying that it appears to represent something that it does not represent. But with this understanding of the idea of a non-veridical concept (which is what really counts, since in discussing my argument what matters is what I mean by it), I think Panaccio’s objections can easily be answered.

In the first place, although Panaccio is right in pointing out that one may be deceived in one’s judgments even with veridical concepts, and even systematically, by an omnipotent deceiver, nevertheless, it should be clear that this is not the only, or even the primary, concern of those who deal with Demon skepticism, especially after Descartes. At any rate, since my argument is clearly not meant to engage any particular historical version of Demon-skepticism, but rather to use an abstract, extreme version of the very idea to bring out the implications of the nominalist conception of concept identity that allows the historical versions to crop up at all, these considerations are not really important. What is important is the question whether on my understanding of non-veridicality the idea of a subject having only non-veridical concepts is really self-contradictory, and thus whether a BIV as I defined it is indeed impossible, and finally, whether the nominalist conception is nevertheless committed to its possibility.

Panaccio’s main objection to my argument showing the impossibility of a BIV as I defined it is that the first premise of my argument that I claim to be self-evident is not self-evident at all, or indeed, it may well be false. However, what we can and what we cannot take to be a self-evident claim clearly depends on its precise understanding; and I think our difference over the self-
evident status of this claim, namely, the claim that it is not possible to form true judgments with non-veridical concepts, hinges on our different understanding of what is involved in my notion of a non-veridical concept. But I hope that the foregoing made it quite clear that when I am talking about a non-veridical concept acquired or planted in our minds under the conditions of complete cognitive isolation from an external reality in a Demon-scenario, then these concepts cannot possibly apply to that reality, and so any judgment meant to apply to that reality, but formed with such concepts, simply cannot be true of that reality. To be sure, this idea may be extreme, but it is not really far-flung from Descartes’ conception of material falsity. So, I think the idea of a BIV as described in my argument is not unjustifiably compared to Descartes’ possibly completely deceived lonely consciousness. But, again, the issue whether my notion of a non-veridical concept is a faithful reconstruction of Descartes’ notion of a materially false idea is beside the point. What matters is that on my understanding of a non-veridical concept, any judgment meant to be about an actual state of affairs of external reality, but formed with non-veridical concepts cannot be true, since its concepts just cannot reach that reality. Or, somewhat more precisely, the semantic values of the terms of the proposition expressing this judgment are simply not elements of the universe of discourse representing that reality. Thus, if I utter the sentence ‘A donkey is brown’, expressing a mental proposition or judgment or thought I form with concepts I acquired in the Matrix, then the supposita of the terms of this sentence as I use it cannot be real donkeys and real brown things, so my judgment simply cannot express the actual state of affairs that some real donkey is identical with some real brown thing. Therefore, this sentence expressing my judgment cannot be true, although, the sentence according to its proper meaning is meant to express that actual state of affairs. So, on this understanding of the semantics of non-veridical concepts, I hope it is clear that the incriminated first premise of my argument must be self-evident, and that it concerns truth and not belief, despite Panaccio’s claim to the contrary. But then, on this understanding, the argument does prove the inconsistency of the notion of a BIV as defined. However, of course, this conclusion is damaging to the nominalist conception, only if it can be shown that the nominalist conception is indeed committed to the possibility of a BIV.

Panaccio does not deny that the nominalist conception is in fact committed to this possibility, at least as a remote, logical possibility, which is all I claimed in my argument. But it is also easy to see why this commitment follows from the nominalist conception, according to which, and this is the crucial point, the veridicality of our simple concepts is their logically contingent feature. For if any and all of our simple concepts is merely contingently veridical, then, since they are veridical or non-veridical logically independently of one another because of their simplicity, it is logically possible that all of them are non-veridical at the same time. But this is precisely the possibility featuring in the definition of a BIV, a cognitive subject whose concepts and generally all mental contents are phenomenally indistinguishable from ours, while all its simple concepts are non-veridical at the same time. So, any thought of such a subject would fail to engage external reality in the semantic sense described above, and so no thought formed by this subject would be true of that reality. Still, this subject is supposed to have the same thoughts as we do, whose thoughts are meant to express actual state of affairs of that reality, such as the thought that this subject is a BIV. But then, this very same thought would have to be true, if this subject were indeed a BIV, and yet it could not be true because, as formed by this subject, it cannot engage this reality. This is what I claim to be the inherent contradiction of the idea of the mere logically contingent veridicality of our simple concepts, which claim, however, is the implication of the nominalist conception of the identity-conditions of concepts.
This much would basically clinch the defense of my argument against Panaccio’s main objection, however, since the point of this discussion is not so much to establish who is right about the consistency of the nominalist position, as to find out what is right to claim about that position, I think it will be instructive to see exactly what it is in the nominalist position that commits it to the logical contingency of the veridicality of our simple concepts, which will then naturally lead to Panaccio’s further interpretive points about the externalism of nominalists as opposed to that of the formal identity theory I attributed to Aquinas.

Another way of characterizing the relevant aspect of the nominalist position in contrast the view I attribute to Aquinas is that the nominalists would accept the possibility that a BIV might have the exact same concepts that we do (assuming now that we are not just brains-in-a-vat), namely, the same concepts, identifiable as such in terms of the same internal properties they have, whereas, while a BIV has these same concepts, it cannot have them about the same things as we do, because by hypothesis, a BIV is isolated precisely from the things we have our concepts about. So, in short, a BIV is supposed to have the same concepts that we do, but cannot have them about the same things we do. However, this means that a conception that allows the possibility of having the very same concepts, but without having them about the same things has to allow the possibility of a BIV. But the nominalist conception clearly allows this possibility, as Panaccio also admitted. Therefore, the nominalist conception is committed to allowing the possibility of a BIV.

But why does this have to be the case with the nominalist conception? And why wouldn’t it be the same on the position I’m attributing to Aquinas? What makes the difference?

The simple possibility I just outlined, namely, possibly having the same concepts but not about the same objects, presupposes that concepts have their conditions of identity at least logically independently from their objects. In other words, the same concepts can be identified in terms of the same internal properties they have (whatever those are; whether a concept is a neural firing pattern of type X, or a spiritual modification of type Y), regardless of what their objects are. Thus, for instance, when I’m thinking of donkeys, I can have neural firing pattern of type X in my brain, while a BIV on this conception could also have firing pattern of type X, which would make it appear to it that it is also thinking of donkeys, but in fact, being a BIV, it cannot possibly think of donkeys, as it was exposed only to virtual donkeys in all its miserable life, in which it never saw a donkey, indeed, it never saw anything. So, what accounts for the emergence of the possibility of Demon skepticism on this account is that concept-identity is construed in terms of a concept’s internal properties (whatever those are), regardless of what the objects of this concept are. In other words, if we identify the objects represented by the concept as its content, i.e., what the concept is about, then we may say that on this conception, the internal properties of the concept do not determine its content, which is precisely the idea that Panaccio hailed as the externalism involved in this conception. And I certainly do not dispute this claim, especially because it evidently supports my claim, namely, that this conception entails the possibility of Demon skepticism as described.

However, is there any other plausible way of construing the identity conditions of concepts? How would the position I’m attributing to Aquinas differ on this issue?

As we have seen, Demon-skepticism in the sense defined is possible, only if our simple cognitive acts are merely contingently veridical, leaving open the possibility that perhaps all our simple cognitive acts are non-veridical. However, if a certain conception of the identity conditions of
these cognitive acts demands that at least some of our cognitive acts are essentially veridical, that is, their veridicality is part and parcel of their conditions of identity, then this conception directly excludes the possibility of Demon-skepticism.

The necessary veridicality of simple cognitive acts with regard to their proper objects is a consequence of the Aristotelian idea that the cognitive act is nothing but the form of the object in the cognitive subject in a different mode of existence. One way of demystifying this apparently obscure description is by appealing to the nowadays common idea of encoding and decoding, i.e., the process of transferring the same information through different media in a way that allows it to be reproducible in a numerically different copy. For instance, the recording and playback of a song is an obvious case of this process. The song played back is a copy of the song originally played, where the reproduction of the song is possible by virtue of the preservation of the same information in the record, which in this sense, is but the form of the song originally played (the modulation of airwaves in the studio) in a different mode of existence, say, existing in the form of the pattern of tiny pits on the surface of a music CD encoding the modulation of airwaves.

Without arguing for it, let us just assume for the time being that this “demystification” correctly captures the original Aristotelian idea. However, even granting this perhaps dubious proposal, one may still have doubts whether it would yield the idea of the necessary veridicality of some simple cognitive acts with regard to their proper objects. After all, just as the pattern of pits on the surface of the CD could in principle be produced by something other than the recording apparatus without the original song actually played in the studio, so the same cognitive act could be produced in the subject without a “matching” object, rendering the act non-veridical, just as the Demon-scenario would suggest. So, apparently, the suggested “demystification” of the Aristotelian idea supports precisely the contingency of the veridicality of cognitive acts and thus the possibility of Demon-skepticism, contrary to what it was devised to illustrate.

However, to proceed from the better known to the lesser known, let us take a closer look at the case of the sound recording. The pattern of tiny pits on the surface of the CD is certainly producible by means other than the recording apparatus. After all, the same kind of laser beam with the same kind of modulation would produce the same pattern, if the modulation of the laser beam were not driven by the modulation of electronic signals driven in turn by the modulation of airwaves hitting the microphones in the recording studio, but, say, by a computer producing the same modulation without any sound whatsoever. However, and this is the important point, in that case the pattern of pits on the surface would not be a record of any sound whatsoever: it may be an ornament, it may be a surface feature, etc., but not a record of some sound. For the pattern of pits to qualify as the record of a song, it has to be part of the system of encoding and preserving information about the actual modulation of air vibrations constituting the song. Indeed, that for the record of a song as such it is essential to encode information about the song whereas it is accidental that it is this pattern of pits in this system of encoding is further confirmed by the fact that if I “rip” the track from the CD onto my computer’s hard drive, then I get the same song onto my hard drive (for if it were not the same song, then the RIAA would certainly have no business harassing me for pirating it), but now recorded in a different medium, this time encoded in the pattern of different magnetic polarities on the surface of the disk.

Describing this process in the language of Aristotelian hylomorphism, we can say that the form of the song that first informed the air in esse reale, existing as the modulation of air waves, first was received in the matter of the CD in esse intentionale, without the matter it originally
informed, merely coinciding with the pattern of pits informing the CD in *esse reale*; then again, it was received in the matter of the hard disk, in another instance of *esse intentionale*, again, without the matter of the original, this time coinciding with the pattern of polarities informing the disk in *esse reale*. Thus, in the whole process, what qualifies any real feature of any medium as the record or encoding of the original form is “the formal unity” of these real features in the sense that the system of encoding secures transferring and preserving the same information throughout the process. If the chain of transferring and preserving the same information is broken, and a merely accidentally similar pattern is produced by some other means, then it may be “misinterpreted” by the next decoder as a recording of some original, but it will never be the same, precisely because it does not fit into the chain in the same way, which is essential for the identity of any encoded bit of information. Thus, to switch to another example, even if a recorded TV program could not be distinguished from the live feed of the same by just looking at the screen, the two are not the same, and their difference is detectable precisely by looking at the process of the transfer of information producing the exact same looking, but essentially different images on the screen.

However, if on the strength of these examples we are willing to interpret the idea of formal unity between cognizer and cognized thing in the sense of the preservation of information, so that this is essential for the identity of the cognitive act insofar as it is an encoding of the form of the object, then it is not hard to see that those simple cognitive acts that are identified precisely in terms of receiving, storing and further processing information about their proper objects will have to be essentially veridical. For then these simple cognitive acts, regardless of what firing patterns of neurons in the brain or what spiritual qualities of an immaterial mind realize them, will only count as the cognitive acts encoding information about their proper objects, if they do in fact represent those objects that they appear to represent to the cognitive subject, for they present or represent to the subject precisely the information they encode about their proper object.

Thus, on this conception, the veridical acts of perception, memory, and intellectual apprehension (as opposed to the non-veridical or contingently veridical acts of hallucination, imagination, misremembering, judging, etc.) are essentially, and not merely contingently veridical. But then, within this conception, the idea of “Demon skepticism” as described earlier is *ab ovo* excluded. Things are as they appear in our veridical acts of cognition, but sometimes, on account of the similarity of a veridical act of cognition to a non-veridical act or to a veridical act of cognizing something else, we may rashly judge things to be the way they appear to be through the non-veridical act, or to be that other thing. But since the veridical act is essentially veridical, and so it cannot be the same as a non-veridical act or the veridical cognition of something else, we can correct our mistake, by detecting the difference, as when we say, “Oh, I thought the bed was on fire, but it was just a dream” or “Oh, I thought I saw water on the road, but it was just a mirage”.

But similar observations apply in the more elaborate cases. For instance, in the scenario of “the Matrix”, the characters eating the peptide goo in physical reality have to realize that when they say it tastes like chicken, they have no genuine conception of chickens, as the only experiences they have about “chickens” are the virtual “chickens” of the Matrix. They could say they had a conception of chickens through those virtual experiences, only if they could look at those virtual experiences as somehow carrying genuine information about genuine chickens, say, if whoever created the program had modeled the virtual chickens after real chickens and presented them as representations of real chickens, in the way a nature video provides us with genuine information.
about genuine animals in remote lands. However if the virtual, quasi-experiences these people had in the Matrix are merely similar to genuine experiences, but are not genuine experiences (whether through direct perception or “mediated perception” as through a documentary), then the concepts abstracted from those quasi-experiences are not the concepts of genuine things that would produce similar, but never the same, experiences. Thus, again, when it comes to the identity conditions of intellectual concepts, which on the Aristotelian account would carry just further processed, abstracted information about the genuine objects of genuine experiences, it is clear that on this conception they also have to be essentially veridical.

So, how is this conception related to the issue of externalism? As we could see, the way this conception identifies concepts has practically nothing to do with their internal properties: we talk about the same concept as long as it is a carrier of the same information whatever realizes it, and what determines this information is precisely the type of external object the concept carries information about. Thus, from the perspective of this conception, whatever internal properties the concept has (say, whether it is a neural firing pattern of a certain type, etc.) is immaterial, since the same concept, carrying the same information, can be realized in just any other type of “medium”. Therefore, the internal properties of the concept not only do not fully determine its content, they have basically nothing to do with it; on this conception the content of the concept is fully externally determined, and so this conception may even be dubbed “hyper-externalism”. At any rate, this is what I meant when I said that Aquinas’ conception involves an even stronger form of externalism than Ockham’s or Buridan’s.

In fact, there is an interesting and important difference between Ockham and Buridan in this regard, which I think sheds some light on their difference from Aquinas as well. When we are talking about the internal properties of a concept that would identify them for a nominalist, we may mean two things: its internal real properties, say, that it is mental act of type X, or its phenomenal representational properties, say, that it is a mental act that makes the cognizer having it aware of things of type Y. At any rate, in Ockham’s case, what Panaccio calls a concept’s “perceptual scheme” and what I would call its “phenomenal content” is clearly different from the concept’s objective semantic content, namely, the objects it actually represents. In Buridan, by contrast, I have seen no evidence of this sort of divergence of phenomenal and semantic content. In fact, Buridan insists that what an absolute concept makes me aware of is just absolutely the object or objects of the type from which I originally acquired my concept. My concept of man makes me aware not of some indifferent, blurry image of human shape, sound and observable behavior (as Ockham’s perceptual scheme or phenomenal content would have it), but rather simply of humans, past, present, future or merely possible (which of course would be the semantic content of this concept for Ockham as well). Still, Buridan finds it logically and hence by divine power perfectly possible to have the same concept without having the same semantic content: God could plant in my mind the same concept, namely, the same mental quality of type X, without however, it representing past, present future, or possible humans. To be sure, this is only a supernatural possibility, but nevertheless it is a possibility for Buridan.

By contrast, for Aquinas, at least on the view I am attributing to him, this is excluded even as a supernatural possibility. God might create a quality in my mind entirely similar in its internal properties to the one whereby I presently conceive of human nature without that quality actually representing human nature, but that quality, not being the encoding of human nature in my mind would not be a concept of humans, just as a pattern of tiny pits on a CD resembling a recording
of a song would not be a recording of the song if it were not the encoding of information about the song. But the same information may certainly be encoded or recorded in different media, yielding again the same representation, realized, however, in entities with radically different properties.

So, on this “hyper-externalist” conception, the formal unity of concept and object, interpreted as the sameness of information encoded in the concept and constituting the object, determines the identity of the concept quite independently from its internal properties. Thus, this “formal unity” does not have to amount to any qualitative similarity between concept and object (or even between my mental act and yours), so of course it is perfectly OK for Aquinas to insist in some passages on the qualitative dissimilarity of concept and object. As he often remarks, what I have in mind when I think of a stone is not the stone, but the species of the stone, and not in the way in which it is in the stone, informing mineral matter, but differently, informing my mind about the form informing that matter. So, the passage Panaccio quotes from Aquinas does not seem to speak against this interpretation of Aquinas as a “conformalist”.

Having thus clarified what I take to be the fundamental differences among the three authors we considered from the point of view their varieties of externalism, let me return in closing to Panaccio’s criticism of my argument, to pinpoint exactly what I could not accept from his interpretation of it. As I noted at the beginning, the main difference between us is in our understanding of what a non-veridical concept is. For Panaccio, a non-veridical concept is simply one with a mismatch between its semantic content and what I would call its phenomenal content, which for Panaccio is just the perceptual “recognition schema” of the concept. Since on Ockham’s view, as Panaccio interprets it, this sort of mismatch is something that may occur even in the ordinary course of nature, of course it is possible to make it systematic and inevitable for an omnipotent deceiver. As I have said, in Buridan I see no evidence for the possibility of this type of divergence, i.e., the natural divergence between phenomenal and semantic content. Indeed, for Buridan, just as for Aquinas, the phenomenal content of an intellectual concept is not anything like Panaccio’s “recognition schema”, because for them, an intellectual concept abstracts precisely from that perceptual scheme, although that scheme may be useful in not infallibly recognizing individuals of the kind represented by the concept. However, for Buridan, there is the possibility of supernatural divergence between phenomenal and semantic content, which would be the case in the Demon-scenario. If it were possible, then in this scenario, if I were a BIV, I would have the very same mental qualities that I now have. And those same mental qualities would appear to me to represent the very same things they do in fact represent now. But then, in that scenario, the same quality just could not latch on to the same objects it is latched on now, since by hypothesis, I am cognitively completely isolated from them. Thus, although I would have the same concepts, I could not have them about the same objects, which is the alleged possibility leading to a contradiction, as I concluded in my argument. But it is precisely this alleged possibility that is correctly deemed to be impossible by Aquinas’ position, at least, as understood in the way I interpret it.
Gyula Klima: Concept Identity and Non-Veridical Concepts

Handout

Concept

1. *Qua entity*, has internal properties (neural firing pattern of type X, spiritual mental act of type Y)
2. *Qua representation*, has content
   a. Phenomenal content: what it makes the subject aware of, what it appears to represent to the cognitive subject
   b. Semantic content: what it carries information about, what it does in fact represent, whether the subject is aware of it or not

*Ockham*: 1 and 2a necessarily go together (but perhaps they are separable by divine power: X may appear once to represent donkeys, once to represent chickens?), whereas 1 and 2b are even naturally separable although they usually go together (concept X that appears to represent eggs and usually does so, may in fact represent marble eggs, say, if my “first oval experience” was an exposure to a marble egg); and they are certainly systematically separable by divine power.

*Buridan*: 1 and 2a and 2b go together by natural necessity, but 1 (“dragging” 2a with itself) is supernaturally separable by divine power, even systematically, allowing the possibility of a BIV scenario.

*Aquinas*: 2a and 2b are *the same*; therefore, they are inseparable by any power: my concept of donkey nature is just donkey nature informing not donkey matter, but my mind, informing my mind *about* the same nature that informs donkey matter. However, 2 is merely contingently related to 1: the same content may be realized in physically rather different mental acts, as the same information may be encoded in different media in different ways.