Contents

Research Articles

Discussion Note

Organon F 31 (2) 2024: 95–113
https://doi.org/10.31577/orgf.2024.31201

ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

An Epistemic-Practical Dilemma for Evidentialism

Byeong D. Lee*

Received: 22 October 2022 / Revised: 29 May 2024 / Accepted: 31 May 2024

Abstract: There are cases in which epistemic rationality seems to conflict with practical rationality. Evidentialists such as Parfit, Shah, Skorupski and Way deny that there are practical reasons for belief. On their view, the only genuine normative reasons for belief are epistemic reasons, and so the alleged practical reasons for belief are the wrong kind of reasons for belief. But I argue in this paper that the evidentialists can still face a genuine dilemma between epistemic and practical rationality which cannot be resolved on the grounds that the alleged practical reasons for belief are the wrong kind of reasons for belief.

Keywords: Epistemic rationality; practical rationality; evidentialism; the right kind of reasons; the wrong kind of reasons.

1. Introductory Remarks

Epistemic rationality is concerned with *what to believe* for the sake of our epistemic goal, which can be understood as having true beliefs (and avoiding false ones). In other words, epistemic rationality is concerned with *what is the case*. By contrast, practical rationality is concerned with *what to do* for realizing what is (practically) desired or desirable. In other words,

* Sungkyunkwan University
https://orcid.org/0000-0001-7080-6702
Department of Philosophy, 25-2, Sungkyunkwan-ro, Jongno-gu, Seoul, Republic of Korea
bydlee@skku.edu

practical rationality is concerned with *what ought to be done*. Clearly, it is one thing to determine what is the case, and it is quite another thing to do what ought to be done. In this sense, epistemic rationality is fundamentally different from practical rationality. But there are cases in which epistemic rationality seems to conflict with practical rationality. One well-known example is this:

> John Doe is suffering from an illness that is usually fatal, but believes with deep conviction that he will recover. The fact that John has this optimistic belief might actually contribute to his recovery. Or at least it might make him more cheerful during his dying days, which in turn might ease the pain of others who are close to him. In either case it would be a virtue or *merit* of John's belief that it has the good consequences that it does for himself or others. (Firth 1998, 259)

Suppose that John's chances of recovery from his illness are 10%. Suppose also that if he believes that he will recover from his illness, his chances of recovery will thereby increase to 30%. Then it seems that John has a good practical (or pragmatic) reason for believing that he will recover.[1] But this optimistic belief will not change the fact that his chances of recovery are still not above 30%. So he has a good epistemic reason for not believing that he will recover. Hence, the John Doe case illustrates a situation in which epistemic rationality seems to conflict with practical rationality. In such a case, what doxastic attitude should John take?

According to pragmatism, there are practical reasons for belief. Contemporary pragmatism is divided into two camps. On the one hand, there is *radical pragmatism*, which holds that, strictly speaking, only pragmatic reasons can be genuine reasons for belief. Versions of this view have been defended notably by Rinard (2015, 2017, 2019a, 2019b) and Maguire and Woods (2020). On the other hand, there is *moderate pragmatism*, which holds that there can be both pragmatic and evidential reasons for belief. Versions of this view have been defended by Foley (1992), Reisner (2008,

---

[1] Note that 'practical reason' is a more comprehensive expression than 'pragmatic reason'. For example, moral reasons, which are distinct from pragmatic or prudential reasons, are also practical reasons rather than epistemic reasons.

2009, 2018), McCormick (2014), Leary (2017), and Howard (2020), among others.

In contrast, evidentialism holds that only evidential reasons can be genuine reasons for belief. This view has been defended notably by Parfit (2011), Shah (2006), Skorupski (2009, 2010), and Way (2012, 2016, 2017). And to reconcile the apparent conflict between epistemic rationality and practical rationality, they distinguish between reasons for belief and reasons for bringing about belief. Let me quote what they say about this distinction:

> But the question *what belief to bring about* is distinct from the question *what to believe*. Answering the former question issues in an action or intention, and thus is determined by practical considerations, such as whether it would be immoral or imprudent to bring about the belief, whereas answering the latter question issues in a belief, and thus is determined by reasons which speak to the truth of the proposition to be believed. (Shah 2006, 498)

> It helps generally distinguish between reasons to ψ and reasons to bring it about that one ψs .... Reasons to believe that your partner is telling the truth are one thing; reasons to make yourself believe it are another. The first are epistemic reasons; the second are practical reasons. You can have practical reason to *make* yourself believe something if you can (e.g., that your partner is telling the truth, or that you will survive the dangerous mission) when there is in fact no reason for you to believe it. (Skorupski 2009, 114–115)

> Since our epistemic reasons are related to the truth of *what* we believe, these reasons can also be called *object-given*. Many people assume that we can also have *state-given* reasons to have certain beliefs. Such reasons would be provided by facts that would make our *having* some belief in some way good. It is often claimed, for example, that we have such reasons to believe that God exists and that we shall have a life after death. These reasons would not be epistemic, or truth-related, but *goodness-related*, or *value-based*. Such alleged reasons to have beliefs are sometimes called *practical* or *pragmatic*. (Parfit 2011, 50–51)

> According to this solution, the right kind of reasons are simply all the reasons there are. Strictly speaking, the wrong kind reasons for attitudes are not reasons for those attitudes, any more than fool's gold is gold. … [I]ncentives for attitudes are not reasons for those attitudes, but are instead reasons to want these attitudes and to bring them about. (Way 2012, 492)

Following Way (2012, 490), we may distinguish between the right and the wrong kind of reasons for belief as follows: The right kind of reasons for believing that *p* are reasons that count in favor of the truth of '*p*'. And the wrong kind of reasons for believing that *p* are considerations that somehow count in favor of believing that *p*, but which don't bear on whether '*p*' is true.[2] On the evidentialist view, evidential reasons for belief are the right kind of reasons for belief, whereas reasons for bringing about belief are the wrong kind of reasons for belief, and the only genuine normative reasons for belief are the right kind of reasons for belief.

With the above view in mind, consider the John Doe case again. In this case, the fact that his optimistic belief about his recovery will increase his chances of recovery significantly is certainly a practical consideration in favor of the belief. But this fact does not give him an epistemic reason for the belief, because his chances of recovery are still not above 30%. As a consequence, on the evidentialist view, there is no real conflict between epistemic and practical rationality for the following reason: John in this case has a good epistemic reason for not believing that he will recover, and he also have a good practical reason for bringing himself to believe that he will recover. But the former is a reason of the right kind for not believing that he will recover, and the latter is a reason of the wrong kind for believing that he will recover.

---

2 More generally, the 'wrong kind of reasons' for an attitude are the kind of reasons that somehow count in favor of the attitude, but which do not bear on the correctness of the attitude. Suppose, for example, that an evil demon will exterminate the human race unless we admire him. In this case, we have a practical reason to admire the demon, but this reason does not bear on whether the demon is really admirable. Reasons of this kind were labeled as the 'wrong kind of reasons' by Rabinowicz and Rønnow-Rasmussen (2004, 393).

I agree with the evidentialists that, strictly speaking, there are no such things as practical reasons for belief. Nevertheless, in this paper, I will argue that the evidentialists can still face a genuine dilemma between epistemic and practical rationality which cannot be resolved on the grounds that the alleged practical reasons for belief are the wrong kind of reasons for belief.

2. The Wrong Kind of Reason for Belief and Epistemic-Practical Dilemma

In his 2016 paper entitled "Two Arguments for Evidentialism", Way provides an argument against reasoning from pragmatic reasons, which he calls "the argument from good reasoning". This argument is roughly as follows:

> Reasons to believe '*p*' must be premises of good reasoning. It is not good reasoning to reason from an incentive for believing '*p*' to believing '*p*'. Therefore, incentives for believing '*p*' are not good reasons to believe '*p*'.

According to Way, reasons are supposed to guide us and the way in which reasons guide us well is through good reasoning; thus, reasons must be premises of good reasoning. In addition, he argues that there is no good reasoning from an incentive for believing '*p*' to believing '*p*'. Consider the following form of argument:

(1) Believing '*p*' is practically beneficial to me. Therefore, '*p*' is true.

One instance of this form of argument is this: Believing that God exists is practically beneficial to me. Therefore, 'God exists' is true. Observe that any argument of this form is not valid. For some false belief could be practically beneficial to someone. So I agree with Way that this kind of argument from an incentive for a belief to the truth of the belief is not good. But it is important to notice at this point that (1) is a theoretical argument rather than a practical argument.[3] And practical reasons must be such that

[3] Following Way (2016, 815), I assume here that good reasoning corresponds to good arguments. What is important to note in this regard is that any piece of good reasoning can be expressed in the form of a good argument.

they can serve as premises of good practical reasoning, rather than premises of good theoretical reasoning. Therefore, showing that there are no genuine practical reasons for belief requires showing that there is no good practical reasoning for belief. Accordingly, we need to think about whether there is no such practical reasoning.

Consider the John Doe case again. The first question worth answering is whether there is an argument to show that John ought to bring himself to believe that he will recover. In this regard, consider the following practical argument:

(2) I ought to promote my survival. Bringing myself to believe that I will recover is the only means of achieving this end. Therefore, I ought to bring myself to believe that I will recover.

Note that (2) is an instance of the following typical form of means-end reasoning:

I ought to achieve end E. My doing A is the only means of achieving E. Therefore, I ought to do A.

This is standardly taken to be a valid rule of practical reasoning. Therefore, insofar as the two premises of (2) are justifiable, (2) can be considered as an argument to show that John ought to bring himself to believe that he will recover. The next question then is whether these two premises are justifiable.

Here we may assume that one's survival is a reasonable goal worth pursuing. So let us move on to the second premise of (2). Is this premise also defensible? What is noteworthy in this context is that the aforementioned evidentialists do not deny that it is possible for one to bring about a certain belief state by some means or other. So let us assume that John can somehow bring it about that he believes that he will recover. Let us also assume that having this optimistic belief is the only possible chance that he has for his recovery. Under these assumptions, the second premise of (2) can also be defended. If so, we have a good practical argument to show that John ought to bring himself to believe that he will recover. The crucial question at this point is whether (2) also provides him with a good reason for believing that he will recover. According to the evidentialists, the answer is 'no'. As mentioned in the previous section, they distinguish between reasons for belief and reasons for bringing about belief, arguing that the latter are the wrong kind of reasons. Therefore, on their view, (2) does not provide John with a reason of the right kind for believing that he will recover.

Unfortunately, however, epistemic rationality can still conflict with practical rationality in a way that cannot be resolved on the grounds that reasons for bringing about belief are the wrong kind of reasons for belief. Consider the John Doe case again. Suppose that his survival is what matters the most to him, and so he wants to do everything he can to increase his chances of recovery. Suppose also that the fact that his chances of recovery improve from 10% to 30% could make a real difference between life and death for him. Under these conditions, it can be argued that John has a good practical reason for bringing himself to believe that he will recover. What is worth recalling at this point is that (2) is a valid practical argument and its two premises could be justified.

Now, suppose that John, as a rational being, wants to comply with what practical rationality demands of him. Then, since he has a good practical reason for bringing himself to believe that he will recover, we can say:

(3) From the practical point of view, John ought to bring it about that he believes that he will recover.

This time, suppose that John, as a rational being, wants to comply with what epistemic rationality demands of him. In this case, his chances of recovery are low, and so he has a good epistemic reason for not believing that he will recover. Accordingly, we may say that he epistemically ought not to believe that he will recover. This is tantamount to saying that it epistemically ought to be the case that he does not believe that he will recover. For this reason, if John comes to believe that he will recover, he can be subject to epistemic criticism on the grounds that this belief violates epistemic rationality. As mentioned in section 1, epistemic rationality is concerned with what to believe for the sake of our epistemic goal of having true beliefs and avoiding false ones. And John's belief that he will recover is likely to be false. Accordingly, what epistemic rationality demands of John is that he should not believe that he will recover. And he fails to meet this epistemic demand if he brings about this belief. Along these lines, we can argue that John can meet what epistemic rationality demands of him only if he does not bring it about that he believes that he will recover. If this is correct, we can also say:

(4) From the epistemic point of view, John ought not to bring it about that he believes that he will recover.

Here two things are worth pointing out. First, even the evidentialists can hardly deny (4). To deny this is tantamount to saying that John is allowed to bring it about that he believes that he will recover. But as argued above, John can meet what epistemic rationality demands of him only if he does not bring about this belief. Second, the reason for John not to bring about this belief is not a reason of the wrong kind in the sense that it is directly relevant to meeting what epistemic rationality demands of him, instead of what practical rationality demands of him. Let me explain. As mentioned in section 1, reasons of the wrong kind for believing that *p* are considerations that somehow count in favor of believing that *p*, but which don't bear on whether '*p*' is true. So we can say that if there is a consideration that counts against believing that *p*, and which doesn't bear on whether '*p*' is true, then that consideration is a reason of the wrong kind for not believing that *p*. But in the John Doe case under consideration, John can meet what epistemic rationality demands of him only if he does not bring it about that he believes that he will recover. In other words, if John brings about this belief, he fails to meet what epistemic rationality demands of him. This implies that whether he brings about this belief is directly relevant to whether he meets what epistemic rationality demands of him. If this is correct, the reason for not bringing about this belief is not a reason of the wrong kind in the sense that it is directly relevant to meeting what epistemic rationality demands of him, instead of what practical rationality demands of him. To put the point another way, the reason for not bringing about this belief bears on whether or not this belief is true. Hence, even the evidentialists can hardly deny (4) just on the grounds that the reason for John not to bring about this belief is a reason of the wrong kind.

In sum, John can meet what practical rationality demands of him by bringing about the optimistic belief, and he also can meet what epistemic rationality demands of him by not bringing about this epistemically unjustified belief. Hence, what epistemic rationality demands of John conflicts with what practical rationality demands of him. What then should he believe or do? Clearly, the aforementioned evidentialist claim, namely that reasons for bringing about belief are the wrong kind of reasons for belief, is not of much help in answering this question. What is worth emphasizing in this regard is that (2) can be a sound practical argument, so that John can have a practical reason of the right kind for bringing about the belief that he will recover. And if he complies with what practical rationality demands of him, then he thereby fails to comply with what epistemic rationality demands of him. One more thing worth pointing out here is that this kind of conflict between epistemic and practical rationality is compatible with the aforementioned reasoning view, according to which reasons, at least primarily, serve as premises of good reasoning.

3. A Genuine Dilemma for Evidentialism

As argued in the previous section, in the John Doe case, John can meet what epistemic rationality demands of him by not bringing it about that he believes that he will recover, and he also can meet what practical rationality demands of him by bringing about this belief. Accordingly, what epistemic rationality demands of John conflicts with what practical rationality demands of him. And as I will argue in the remainder of this paper, the evidentialists cannot resolve this dilemma in a principled way.

Suppose, *for reductio*, that they could resolve this dilemma in a principled way. Then there are two possibilities. One possibility is that it is the right thing for John not to bring himself to believe that he will recover, all things considered. The other possibility is that it is the right thing for John to bring about this belief, all things considered.

Let us begin by considering the first possibility. To say that this possibility is the case is tantamount to saying that epistemic rationality overrides practical rationality when they conflict with each other. Then we can make the following claim:

(5) John's epistemic reason for not believing that he will recover overrides his practical reason for bringing about this belief.

But the evidentialists can hardly defend this claim. Let me explain.

To begin with, as pointed out in section 1, epistemic rationality is concerned with believing what is true, whereas practical rationality is concerned with bringing about what is (practically) desired or desirable. Clearly, it is one thing to determine *what is the case*, and it is quite another thing to do *what ought to be done*. In this sense, epistemic rationality is fundamentally different from practical rationality.

In addition, if John believes that he will recover, then this belief plays dual roles. On the one hand, it is a belief whose content can be either true or false. On the other hand, it is also a condition necessary for promoting his survival, which is a practical goal that he can hardly give up. And this latter role requires John to bring about a state necessary for promoting his survival, and this state happens to be a certain belief state. So it is a contingent fact about John that this belief plays these dual roles. And epistemic rationality is concerned with the first role, and practical rationality is concerned with the second role.

Moreover, an exceptional case like the John Doe case is not a reason to revise our present epistemic or practical norms. Let me elaborate on this point a bit further.

As pointed out before, it is one thing to make an epistemic evaluation of whether or not a certain belief is true, and it is another thing to make a practical evaluation of whether a certain action is required for realizing what is desired or desirable. And we make such an epistemic evaluation in terms of our epistemic norms, and such a practical evaluation in terms of our practical norms. Another important thing to note is that our epistemic norms are primarily *intersubjective* norms, which have normative force for us in our social practice of demanding justification and responding to such demands. For example, our beliefs are bound by modus ponens. So, if you believe not only that if p then q, but also that p, and if you care whether q, then you ought to believe that q. Of course, someone can believe in a way that violates some epistemic norms such as modus ponens. But such a person can be subject to rational criticism on the grounds that he or she violates an epistemic norm of rationality. A similar point can be made about our practical norms. Our practical norms are primarily *intersubjective* norms, which have normative force for us in our social practice of justification. For example, our actions are bound by the following means-end reasoning: If you ought to achieve end *E*, and if doing *A* is a means implied by your achievement of *E*, then you ought to do *A*. Again, someone can act in a way that violates some practical norms such as means-end reasoning. But such a person can be subject to rational criticism on the grounds that he or she violates a practical norm of rationality.

As mentioned before, John's belief that he will recover plays dual roles: a belief whose content can be either true or false, and a condition necessary for promoting his survival. And epistemic rationality is concerned with the first role, and practical rationality is concerned with the second role. Accordingly, the first role is evaluated in terms of our epistemic norms, and the second role is evaluated in terms of our practical norms. Now observe that, on the basis of our present epistemic norms, we can judge that John epistemically ought not to bring it about that he believes that he will recover, and also that, on the basis of our present practical norms, we can judge that he practically ought to bring about this belief. At this point, three things are worth pointing out. First, the former judgment is concerned with the aforementioned first role of John's belief, and the latter judgment is concerned with the aforementioned second role of this belief. Second, there is nothing wrong with these judgments. Third, an optimistic but improbable belief could be practically beneficial to someone, and there is no mystery about this familiar fact. Therefore, an exceptional case like the John Doe case is not a reason to revise our present epistemic or practical norms.

In addition to the above reasons, the view that epistemic rationality overrides practical rationality has a very implausible consequence. If this view is correct, then it ought to be the case that John does not believe that he will recover, all things considered. Then it should be the case that the evidentialists can rationally demand (or advise) of John that he not bring it about that he believes that he will recover. But this is tantamount to demanding that he give up his efforts for promoting his survival by holding the optimistic belief that he will recover. And it can be argued that nobody has the right to demand that one give up one's efforts to survive, at least insofar as those efforts do not infringe anyone's rights.

Suppose that John's survival is what matters the most to him, and so he is willing to do everything he can to survive. Suppose also that his optimistic belief that he will recover could make a crucial difference for his survival. To put it another way, the fact that his chances of recovery improve from 10% to 30% could make a real difference between life and death for him. As a consequence, if John survived his illness (because of this optimistic belief), he could argue that, if he did not hold this belief, he could not have survived. Suppose further that John could somehow bring about this optimistic belief. Under these conditions, (2) can be a sound argument for John.

(2) I ought to promote my survival. Bringing myself to believe that I will recover is the only means of achieving this end. Therefore, I ought to bring myself to believe that I will recover.

Let us also assume that John's efforts to survive do not infringe anyone's rights. In such a case, we can say that it is practically rational for John to bring it about that he believes that he will recover. And we can hardly argue that he ought to give up his efforts to survive, just on the grounds that the chances of his recovery are low. What is worth considering in this context is Kant's claim that practical reason has primacy over theoretical reason. He says:

> Thus, in the union of pure speculative with pure practical reason in one cognition, the latter has primacy, assuming that this union is not *contingent* and discretionary but based a priori on reason itself and therefore *necessary*. … But one cannot require pure practical reason to be subordinate to speculate reason and so reverse the order, since all interest is ultimately practical and even that of speculative reason is only conditional and is complete in practical use alone. (Kant 1996, 5:121)

> [I]n the end all the effort of our faculties is directed to what is practical and must be united in it as their goal. (Kant 2000, 5:206)

On Kant's view, thus, all interest is ultimately practical, and in the end all the effort of our faculties is directed to what is practical. Therefore, insofar as theoretical reason has an interest, this interest must be ultimately practical. Let me elaborate on this point a bit further.

As mentioned before, epistemic rationality is concerned with our epistemic goal of having true beliefs (and avoiding false ones). So we may say that the interest of theoretical reason consists in having true beliefs. But one important question arises here: why should we be interested in having true beliefs in the first place? One plausible answer might be that we need true beliefs about the world in order to make rational decisions for our ultimate practical goal, such as our survival and well-being or the Kingdom of Ends. Note that all living animals need correct information about the world necessary for their survival and well-being. We are no exception. One important difference between mere animals and us is that we can engage in theoretical reasoning in order to obtain true beliefs about the world. In other words, unlike mere animals, we are rational beings whose beliefs are bound by epistemic norms of rationality. But notice that such epistemic norms would be pointless to us, if they are of no use for us to make rational decisions for our ultimate practical goal. To put the point another way, we can hardly enforce such epistemic norms on people if those norms are useless or even detrimental for their survival and well-being. Along these lines, we may argue that the reason why we should be interested in theoretical reason (or our epistemic goal of having true beliefs and avoiding false ones) is that we need true beliefs about the world in order to make rational decisions for our ultimate practical goal. It is in this sense that practical reason has primacy over theoretical reason. If these considerations are on the right lines, it is very unlikely that epistemic rationality overrides practical rationality when they conflict with each other.

Let us now turn to the second possibility, namely, that it is the right thing for John to bring about the belief that he will recover, all things considered. If this possibility is the case, the evidentialists can rationally demand (or advise) of John that he bring it about that he believes that he will recover. As I argue below, however, there are important reasons to think that they also can hardly make this demand.

The first reason is concerned with the *ought-implies-can* principle. As Kant (1998, A807/B835) insists, ‘ought’ implies ‘can’. The most important theoretical rationale for this ought-implies-can principle is that it is pointless to demand of any person that they do what they are unable to do. The question then is whether John can bring it about the belief that he will recover. Recall that the chances of his recovery are low. And insofar as he is well aware of this fact, he can hardly believe that he will recover. In this context, it is important to note that beliefs are not actions but *mental states*. An action is essentially something one can do on purpose. In contrast, a mental state is what occurs in a person's mind, rather than something one does on purpose. So, although you may decide to perform a certain action, you cannot directly decide to make a mental state occur in your mind. In other words, it is not under your direct volitional control to make a certain mental state occur in your mind.[4] For this reason, if you believe that there are good reasons for the truth of '*p*', then, under normal circumstances, you are thereby disposed to believe that *p*. For example, if you believe that '*p*' follows by modus ponens from things you firmly believe, you are thereby disposed to believe that *p*. Therefore, at least under normal circumstances, we do not form beliefs directly by willing to believe them. This is why John can hardly believe that he will recover just by intending to believe it, especially in the face of strong evidence against it. Hence, in order to bring about the belief that he will recover, he needs to ignore evidence against this belief. To put the point another way, he needs to engage in self-deception on this matter. But the problem is that engaging in self-deception is not something that one can normally do on purpose. What is noteworthy in this regard is that John is confronted with a dilemma, when he intends to ignore evidence against his belief that he will recover. Insofar as he does not forget that he needs to engage in self-deception so as to form and maintain this optimistic belief, he is (implicitly) aware that he is unlikely to recover. Under this condition, he can hardly succeed in really believing that he will recover. If he somehow forgets that he is engaging in self-deception so as to maintain this belief, then he is very likely to lose this belief as soon as he is again confronted with the compelling evidence against it. Therefore, self-deception of this sort tends to be very unstable. If, however, engaging in self-deception is not something that John can normally do on purpose, the evidentialists can hardly demand that he bring about the belief that he will recover through self-deception. To put the point another way, John can refuse this demand on the grounds that he can hardly meet this demand.

[4] For a more detailed discussion of this point, see Sellars, 1967, esp. 74.

The second and related reason against the second possibility is that if one holds a false belief by engaging in self-deception, then one can be more vulnerable to making wrong moral or other decisions on the basis of this false belief. Let me explain.

There are cases in which deceiving a person could bring about good results for him. But this does not show that in such a case it is rationally permissible to deceive him. It is important to observe at this point that what one believes can serve as reasons for moral or other decisions. And we can hardly rule out the possibility that the deceived person could make a wrong decision on the basis of the false belief. In a similar vein, there are cases in which engaging in self-deception could bring about good results. But this does not show that in such a case it is rationally permissible to deceive oneself. For one thing, if one holds a false belief through self-deception, then one can be more vulnerable to making wrong moral or other decisions on the basis of this false belief. For example, if John believes that he will recover through self-deception, he may miss the opportunity to address personal matters before his potential passing, such as settling debts and ensuring the well-being of his children in his absence. Besides, if he will not recover, perhaps he would be better off accepting this fact and spending his remaining days as meaningful as he can.

Now suppose that John refuses to engage in self-deception because he does not want to compromise his own moral or intellectual integrity, and also because he wants to spend his remaining days as meaningful as he can, while fully understanding his real situation. In such a case, we can hardly say that he deserves criticism or blame. Especially, from the evidentialist perspective, there is nothing wrong with John's decision not to bring about the belief that he will recover. If this is correct, the evidentialists can hardly argue that John ought, all things considered, to bring it about that he believes that he will recover.

There is one more thing worth considering here. As pointed out before, practical reason has primacy over theoretical reason in the sense that we need true beliefs in order to make rational decisions for our ultimate practical goal. But the primacy of practical reason in this sense does not imply that whether or not a belief is true is affected by our ultimate practical goal. Insofar as we need to make rational decisions on the basis of true beliefs, we have to evaluate whether or not a relevant belief is true on the basis of our epistemic norms. And as emphasized before, it is one thing to determine on the basis of our epistemic norms whether a belief is true, and it is quite another thing to determine on the basis of practical norms whether one ought to realize a certain state of affairs. Furthermore, recall that the evidentialists deny that there are practical reasons for belief. Thus, our epistemic judgment of whether '*p*' is true or not is not the kind of thing which can be defeated by any practical consideration. For example, the fact that John Doe has a practical reason for bringing about the belief that he will recover has no effect on the fact that he is not epistemically justified in holding the belief. Therefore, although practical reason has primacy over theoretical reason in the sense that we need true beliefs in order to make rational decisions for our ultimate practical goal, epistemic reasons for belief are not such that they can be defeated by any practical reasons.

In sum, the evidentialists can hardly demand that John not bring it about that he believes that he will recover, because nobody has the right to demand that one give up one's efforts to survive, at least insofar as those efforts do not infringe anyone's rights. In addition, they can hardly demand that John bring himself to believe that he will recover, either, because he can rightly refuse this demand on the grounds that this belief is not epistemically justified and he wants to spend his remaining days as meaningfully as he can, while fully understanding his real situation. Hence, the evidentialists cannot resolve this kind of epistemic-practical dilemma in a principled way.

4. Concluding Remarks

The evidentialists deny that there are practical reasons for belief. On their view, the only genuine normative reasons for belief are epistemic reasons, and so the alleged practical reasons for belief are the wrong kind of reasons for belief. But if the arguments presented in this paper hold, they can still face a genuine dilemma between epistemic and practical rationality which cannot be resolved on the grounds that the alleged practical reasons for belief are the wrong kind of reasons for belief.

I argued for the above claim by focusing on the John Doe case. In this case, John's belief that he will recover is likely to be false, and so it is not epistemically justified. Therefore, we may say that from the epistemic point of view, John ought not to believe that he will recover. This, in turn, requires him not to bring it about that he believes that he will recover. The reason is clear. If he does so, he thereby fails to comply with what epistemic rationality demands of him. By contrast, if John believes that he will recover, this optimistic belief increases his chances of recovery significantly. Therefore, we may also say that from the practical point of view, John ought to bring about this belief. Hence, what epistemic rationality demands of John conflicts with what practical rationality demands of him. And the evidentialists cannot resolve this conflict in a principled way.

To begin, the evidentialists can hardly demand that John not bring it about that he believes that he will recover. To demand that he not do so is tantamount to demanding that he give up his efforts for promoting his survival by holding the optimistic belief that he will recover. But nobody has the right to demand that one give up one's efforts to survive, at least insofar as those efforts do not infringe anyone's rights. And one's right to survive is not overridden by the claim that reasons for bringing about belief are the wrong kind of reasons for belief.

In addition, the evidentialists can hardly demand that John bring it about that he believes that he will recover, either. Recall that John can rightly refuse this demand on the grounds that this belief is not epistemically justified, and he wants to spend his remaining days as meaningfully as he can, while fully understanding his real situation. Recall also that, at least from the evidentialist perspective, epistemic rationality is not overridden by practical rationality, and so there is nothing wrong with this refusal.

If the above considerations are correct, the evidentialists can still face a genuine dilemma between epistemic and practical rationality which cannot be resolved on the grounds that the alleged practical reasons for belief are the wrong kind of reasons for belief. And insofar as they are right about the claim that epistemic rationality is not overridden by practical rationality, all of us can face such a genuine dilemma which cannot be resolved in a principled way.

References

Firth, Roderick. 1998. "Epistemic Merits, Intrinsic and Instrumental." *In Defense of Radical Empiricism: Essays and Lectures*, edited by John Troyer, 259-71. Lanham: Rowman & Littlefield. https://doi.org/10.2307/2693706

Foley, Richard. 1992. *Working without a Net: A Study of Egocentric Epistemology*. Oxford: Oxford University Press. https://doi.org/10.1111/0029-4624.00100

Howard, Christopher. 2020. "Weighing Epistemic and Practical Reasons for Belief." *Philosophical Studies*, 117(8): 2227-2243. https://doi.org/10.1007/s11098-019-01307-y

Kant, Immanuel. 1996. "Critique of Practical Reason." In *Practical Philosophy*, translated and edited by Mary J. Gregor, 133–271. Cambridge: Cambridge University Press.

Kant, Immanuel. 2000. *Critique of the Power of Judgment* (*The Cambridge Edition of the Works of Immanuel Kant*), edited and translated by Paul Guyer, translated by Eric Matthews. Cambridge: Cambridge University Press.

Leary, Stephanie. 2017. "In Defense of Practical Reasons for Belief." *Australasian Journal of Philosophy*, 95(3): 529-542. https://doi.org/10.1080/00048402.2016.1237532

McCormick, Miriam Schleifer. 2014. *Believing against the Evidence: Agency and the Ethics of Belief*. New York: Routledge. https://doi.org/10.4324/9780203579146

Maguire, Barry, and Jack Woods. 2020. "The Game of Belief." *Philosophical Review*, 129(2): 211-249. https://doi.org/10.1215/00318108-8012843

Parfit, Derek. 2011. *On What Matters. Volume One*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199572809.001.0001

Reisner, Andrew. 2008. "Weighing Pragmatic and Evidential Reasons for Belief." *Philosophical Studies*, 138(1): 17-27. https://doi.org/10.1007/s11098-006-0007-3

Reisner, Andrew. 2009. "The Possibility of Pragmatic Reasons for Belief and the Wrong Kind of Reasons Problem." *Philosophical Studies*, 145(2): 257-272. https://doi.org/10.1007/s11098-008-9222-4

Reisner, Andrew. 2018. "Pragmatic Reasons for Belief." In *The Oxford Handbook of Reasons and Normativity*, edited by Daniel Star, 705-728. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199657889.013.31

Rinard, Susanna. 2015. "Against the New Evidentialists." *Philosophical Issues*, 25(1): 208-223. https://doi.org/10.1111/phis.12061

Rinard, Susanna. 2017. "No Exception for Belief." *Philosophy and Phenomenological Research*, 94(1): 121-143. https://doi.org/10.1111/phpr.12229

Rinard, Susanna. 2019a. "Believing for Practical Reasons." *Noûs*, 53(4): 763-784. https://doi.org/10.1111/nous.12253

Rinard, Susanna. 2019b. "Equal Treatment for Belief." *Philosophical Studies*, 176(7): 1923-1950. https://doi.org/10.1007/s11098-018-1104-9

Rabinowicz, Wlodek, and Toni Rønnow-Rasmussen. 2004. "The Strike of the Demon: On Fitting Pro-attitudes and Value." *Ethics*, 114: 391-423. https://doi.org/10.1086/381694

Sellars, Wilfrid. 1967. *Science and Metaphysics: Variations on Kantian Themes*. London: Routledge & Kegan Paul.

Shah, Nishi. 2006. "A New argument for Evidentialism." *Philosophical Quarterly*, 56(225): 481-498. https://doi.org/10.1111/j.1467-9213.2006.454.x

Skorupski, John. 2009. "The Unity and Diversity of Reasons." In *Species of Reason: New Essays in the Philosophy of Normativity*, edited by Simon Robertson, 113-139. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199572939.003.0005

Skorupski, John. 2010. *The Domain of Reasons*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199587636.001.0001

Way, Jonathan. 2012. "Transmission and the Wrong Kind of Reason." *Ethics*, 122: 489-515. https://doi.org/10.1086/664749

Way, Jonathan. 2016. "Two Arguments for Evidentialism." *The Philosophical Quarterly*, 66(265): 805-818. https://doi.org/10.1093/pq/pqw026

Way, Jonathan. 2017. "Reasons as Premises of Good Reasoning." *Pacific Philosophical Quarterly*, 98(2): 251-270. https://doi.org/10.1111/papq.12135

Organon F 31 (2) 2024: 114–140
https://doi.org/10.31577/orgf.2024.31202

ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Riemann's Philosophy of Geometry and Kant's Pure Intuition

Dinçer Çevik*

Received: 1 January 2022 / Revised: 15 March 2024 / Accepted: 9 April 2024

Abstract: The aim of this paper is twofold: first to explicate how Riemann's philosophy of geometry is organized around the concept of manifold. Second, to argue that Riemann's philosophy of geometry does not dismiss Kant's spatial intuition. To this end, first I analyse Riemann's *Habilitationsvortrag* with respect to interaction between philosophical, mathematical and physical perspectives. Then I will argue that although Riemann had no particular commitment to the truth of Euclidean geometry his alternative geometry does not necessarily dismiss Kant's spatial intuition.

Keywords: G.F.B. Riemann; Kant; pure intuition; non-Euclidean geometries.

1. Introduction

Although Georg Friedrich Bernhard Riemann's greatness in mathematics has been well acknowledged, and the importance and implications of his geometry studied widely by philosophers, the same does not seem to be true of his philosophy of geometry. In part, this paper is motivated by this very fact.

* Mugla Sitki Kocman University
https://orcid.org/0000-0001-5897-7381
Mugla Sitki Kocman University, Faculty of Letters, Department of Philosophy, Mugla, 48000, Turkey
dincercevik@mu.edu.tr

In his *Habilitationsvortrag* of 1854, Riemann sets aside the usual approaches that had been taken until then, and instead tries out new ideas and approaches. Riemann wanted to depict nature from the perspective of its inner structures and one aspect of this endeavour entailed questioning the nature of space and geometry from heterogeneous points of view, such as mathematics, physics, and philosophy.[1] Riemann thought that while Euclidean geometry made an interesting proposal for the construction of a theory of space, there was in fact no *a priori* connection between the concept of space and the axioms of Euclidean geometry. He argued, then, that the fundamental concepts central to Euclidean geometry do not have to be part of every system of geometry imaginable. That is, the fundamental concepts of Euclidean geometry should not be thought of as necessary for the construction of all possible systems of geometry. In order to reach these conclusions about Euclidean geometry, and in order to introduce new concepts, it was necessary for Riemann to engage in the activity of conceptual clarification. The fundamental new concept he introduced was the concept of manifold. Describing this notion, Riemann explicitly refers to Johann Fredrich Herbart and Carl Fredrich Gauss. Since Herbart and Gauss were very critical of Kant's philosophy of geometry, Riemann—under their influence—also makes critical remarks about Kant's philosophy of geometry. However, in this paper I will argue that Riemann's alternative geometry does not necessarily dismiss Kant's spatial intuition.

2. The Architecture of Habilitationsvortrag of 1854

Riemann discusses the problem of what he calls 'multiply extended magnitude' in his famous lecture 'On the Hypotheses Which Lie at the

[1] Riemann gives a hint about his research project in an undated note 'My principal task is a new interpretation of the well laws of nature-their expression by means of other fundamental concepts-that would make possible the utilization of experimental data on the interaction of heat, light, magnetism, and electiricity for the investigation of their correlations'. For an exposition of Riemann's philosophy of nature see Bottazzini & Tazzioli (1995).

Foundation of Geometry'. Riemann's introduction clearly shows that he saw himself involved in a philosophical as well as mathematical enterprise:

> It is well known that geometry presupposes not only the concept of space but also the first fundamental notions for constructions in space as given in advance. It gives only nominal definitions for them, while the essential means of determining them appear in the form of axioms. The relation of these presuppositions is left in the dark; one sees neither whether nor in how far their connection is necessary, nor a priori whether it is possible. (1929/1959, 411)

Riemann claims that throughout history neither mathematicians nor philosophers shed light on the 'darkness' that lies at the foundations of geometry. In this regard, he thought that the reasons for this ambiguity lied in the fact that the general concept of multiply extended magnitudes had not been investigated, and that the ideas that properties depend on shape and that metrical properties depend on measure had not yet been properly separated. Accordingly, Riemann set himself two tasks: The first (a philosophical task) was to define a manifold extension. The second (an empirical task) was to give definitions of intrinsic curvature and measure determined from within extension. For the second he was indebted to Gauss, and the first to Herbart. Riemann neatly divides *Habilatitionsvortrag* into three parts, and as such one must analyse it in accordance with its philosophical, mathematical, and physical characteristics.

2.1. Philosophy in Habilitationsvortrag

Riemann introduced certain new and fruitful concepts into the discussion about geometry and space. For example, the discrete and continuous structure of space and the problem of measurement related to it; intrinsic features of space (that is those features that could only be determined without considering the fact that space is embedded in a higher-dimensional space) and extrinsic features (that are properties of this embedding); the problem of metric; and intrinsic and extrinsic metric. All of these new concepts imply a new vision for geometry; however, the concept of manifold stands at the centre of Riemann's new understanding of the subject.[2] By means of this notion, Riemann developed a new ontology of space. His concept of manifold pre-exists in an epistemic sense and it is logically prior to the concept of space (Gray in Laugwitz 1999, 235; Scholz 1992, 23). Riemann's main concern was construction *of* space, rather than construction *in* space. Riemann reasons as follows: Take a concept from any field of investigation, then think of a concept 'whose mode of determination varies continuously'; if one proceeds in 'a well determined way' from one mode to another, one gets a simply extended manifold. If one proceeds to pass over from each point of a manifold to another this procedure will result in two-dimensional (doubly extended) manifold. If we continue this procedure from two-dimensional manifold to another we will get to a triply extended manifold. Here it is important to note that in the one-dimensional case we can only move in one direction: forwards and backwards. So, in order to define motion on two-dimensional manifolds (i.e. surfaces), we have to speak of two different directions; in the case of three-dimensional mani-

[2] Before Riemann, Kant had also used the concept of manifold in *Prolegomena*, *Metaphysical Foundations of Natural Science,* and *Critique of Pure Reason*. In this sense it can be argued that Riemann owes this concept to Kant (Plotnitsky 2009, 112). However, Ferreiros (2004, 4) argues that Riemann owes the term manifold to his teacher Gauss. I agree with Ferreiros' interpretatation. In the beginning of Habilatitionsvortrag, Riemann refers to Gauss' studies of biquadratic residues in the 1832 announcement of that paper, and his 1849 proof of the fundamental theorem of algebra. All of these works are related to complex numbers (Nowak 1989, 27, Ferreiros 2007, 44). Gauss speaks of 'a manifold of two dimensions' in his interpretation of complex numbers. Gauss understands 'manifold' as nothing but system of objects connected with relations. These relations have some interconnections and properties that determine the dimensionality of manifold. Hence, Gauss wants to pay attention to properties by means of which it would be possible to consider a physical system as a two-dimensional manifold (Ferreiros 2007, 44). Gauss makes use of geometric language in a non-geometric context. Separating possibility of mathematics based on abstract spatial concepts from a constrained approach derived from perception, he discusses the geometry of the complex numbers (Nowak 1989, 27). More importantly, he talks about continua of n-tuples of numbers. He takes points of a plane determined by the coordinates t, u, and introduced an algebraic structure of complex numbers. Similarly, Riemann was to introduce real n-tuples and to investigate a 'metric structure' (Laugwitz 1999, 226).

fold (space), three different directions. n-dimensional manifolds can be understood in a similar way (i.e., where we can move in n different directions). So, we can simply say that a manifold is constructed in the relation between 'a variable object' and its capability of taking different states ('forms or modes of determinations'); these different states comprise the 'points' of manifold.[3] Riemann coined the term "Mannigfaltigkeit" to describe a set where "Bestimmungsweisen" ("ways of determination") constitute its instances or specializations. There are various interpretations of Riemann's concept. For instance, one could consider the permitted singularities within a "Mannigfaltigkeit". Riemann illustrated his idea using the set of colours, suggesting it possessed three dimensions. Additionally, he referred to a "Riemann" surface as a mathematical example.

In contemporary language, concept of manifold should be understood as a set characterized by n -tuples of real numbers. However, no formal definition is provided at the beginning of the *Habilitationsvortrag* (Ohshika 2017, 300). Hermann Weyl provides examples of manifolds, illustrating that the distinct conditions of equilibrium of an ideal gas, characterized by two independent variables like pressure and temperature, constitute a two-dimensional manifold. Similarly, the points on a sphere, or the system of pure tones described in terms of intensity and pitch, represent other examples. Additionally, based on physiological theory, which posits that colour sensation is influenced by three chemical processes occurring on the retina—black-white, red-green, and yellow-blue—each with specific directions and intensities, colours form a three-dimensional manifold in terms of quality

[3] Ohshika (2017, 295) underlines that the concept of manifold stands as a crucial cornerstone in modern geometry, and even in contemporary mathematics as a whole. Its inception is commonly attributed to Riemann, with the term "Mannigfaltigkeit," translated into English as "manifold," making its debut in Riemann's renowned *Habilitationsvortrag*. He also explains that while "manifold" is the most common translation, alternative English renderings like "multiplicity" or "variety" can be found in literature. Furthermore, he adds that, "Mannigfaltigkeit" had been used in non-mathematical contexts prior to Riemann's work, including a poem by Schiller titled "Mannigfaltigkeit." To explore the inception, elaboration, and evolution of the manifold concept, beginning with Riemann's *Habilitationsvortrag*, refer to (Ohshika, 2017). Alternatively, for an examination of the historical trajectory of the manifold concept from Grassmann through Riemann to Husserl, see Morales (2019).

and intensity. However, colour qualities alone form a two-dimensional manifold. Weyl highlights that the defining feature of an n-dimensional manifold is that every element within it (whether individual points, gas conditions, colours, or tones) can be precisely described by providing n quantities, referred to as "coordinates," which are continuous functions within the manifold (1922, 84). The concept of manifolds hints at the prospect of conceptualizing and potentially defining a space based on its relationships with other spaces (Plotnisky 2017, 350). In the definition of an n-fold extended manifold it is crucial to note that n represents the count of independent directions available for movement and that manifold concept is related with local uniqueness of the way connecting two points.

Riemann's philosophical concerns had already appeared before *Habilitationsvortrag*, when he felt the need to outline a new approach to geometry (Scholz, 1982). This new approach was philosophical in character and Herbartian in spirit. Although there is little agreement concerning exactly to what extent Herbart influenced Riemann, in its main aspects Riemann's view of mathematics benefits from a comparison with certain points of Herbart's philosophy[4]. Riemann's published works contain philosophical fragments that shed some light upon his reflections about science and which provide evidence that Riemann was strongly influenced by Herbart. Specifically, Erhard Scholz's (1982) essay contains extracts from Riemann's *Nachlass* that indicate that mathematics from Riemann's point of view and philosophy as seen by Herbart share some fundamental similarities. Riemann's selections of passages from texts by Herbart suggest that he was particularly interested in the problem of change and the structure of reality. Herbart differentiates appearance and reality. In Herbart's view, experience shows us properties and bundles of properties, while the underlying reality must be searched for within the things to which properties are ascribed. This distinction between the phenomena and a more stable underlying reality, and an investigation of the relationship between them, is essential in Riemann's own reflections about the epistemology of science. Based on Herbart's distinction between changing phenomena and underlying reality, Riemann constructs his methodology of science. However, Herbart's idea of

[4] See Russell (1956), Torretti (1978), Scholz (1982), Ferreiros (2007), Banks (2005), and Werner (2010).

the advancement of knowledge is modified by Riemann. While Herbart seems to explain the process of knowledge in metaphysical way, Riemann's modified view is closer to a form of scientific research.

2.2. Herbart on Space(s), Avoiding a priorism, and Orientation of Mathematical Research

Herbart treats the Kantian understanding of space and time as 'a completely shallow, meaningless, and inappropriate [*völlig gehaltlose, nichtssagende, unpassende*] hypothesis'; that is, as naming them 'inborn' and 'empty containers' (Scholz 1982, 421). According to Herbart, spatial concepts are no different from all other concepts, which serve as 'forms of experience'.[5] Like all concepts, the origin of spatial concepts is found in experience. Yet, through philosophical and scientific thinking we give to shape spatial concepts. Space and time are departures from which Herbart produces more broad 'continuous serial forms' (*continuierliche Reihenformen*). Herbart sees things as 'bundles of properties', so for him any property can be considered a 'qualitative continuum'. Thus, for Herbart continuous serial forms mean a pure flux of instantaneous, space-less sensations that undergo dynamical, reciprocal changes among themselves. Herbart's main examples are the 'line of sound' and a coloured triangle with blue, red, and the yellow at the corners and mixed colours in the two-dimensional continuum in between.[6] The basic idea of continuous serial forms is to transfer spatial concepts into a non-geometric context. It seems likely that Herbart's theory of Serial forms (*Reihenformen*) was stimulating for Riemann, and played a role in the formation of the concept of manifold. In *Habilitationsvortrag*, Riemann, mentioning colour when talking about continuous positions in space, and using the word 'transition' between modes of determination when introducing concept of manifold, evokes a Herbartian

[5] In this sense it can be argued that for Herbart spatial concepts are like Kant's space and time as pure forms of intuition.

[6] In this context, it is possible to refer to some important works on a "space of colour" accomplished during the 18th and 19th centuries by Johann Wolfgang von Goethe and Philipp Otto Runge. See Barsan & Merticariu (2016).

construction of extended magnitude by means of continuous transitions between qualities (Banks 2005, 228; Scholz 1982, 422).[7]

Based on his theory of psychological space, Herbart wanted to discuss space with respect to the difference between *intelligible* and *phenomenal* space. The quote from Herbart below is very similar to something Riemann says in the beginning of the *Habilitationsvortrag*:

> Geometry assumes space as given; and it makes its constituents, lines and angles, through construction. But for the simple essences (and natural philosophy must be reduced to them in order to find solid ground of the real) no space is given. It together with all its determinations must be produced. The standpoint of geometry is too low for metaphysics. Metaphysics must first make clear the possibility and validity of geometry before she can make use of it. This transpires in the construction of intelligible space. (Herbart in Lenoir 2006, 152)

[7] Riemann prefers continuous manifolds over discrete manifolds. About the reasons for this preference see Laugwitz (1999, 307–308), who argues that continuous manifolds derive their existential quality from the realm of the conceptual. On the other hand, Ferreiros (2007, 58) holds the view that continuous manifolds are firm basis for the generalization of Gauss' differential geometry. Riemann's preference also seems to be compatible with an understanding of Herbart's philosophical speculations. Riemann seems to suggest a Herbartian construction of extended magnitude by means of a continuous transition between qualities (Banks 2005, 228). According to Scholz (1992, 22), since analyzing the concept of the continuity came after the emergence of formal definitions of real numbers and the formulation of set theoretical ideas, his preference must be interpreted in an 'intuitive sense'. Russell (1956, 14) on the contrary claims that Riemann prefers the discrete above the continuous. In opposition to Russel's claim, Torretti (1978, 108) argues that 'I do not know what Russell had in mind when he spoke of "Herbart's his general preference for the discrete above the continuous", so that I cannot judge wherein such preference shows up in Riemann's writings'. I think Torretti is right on this since we see clear evidence when Riemann explicitly stresses the importance of the continuous over the discrete. In *Habilatitionsvortrag* there are a number of places in which Riemann stresses this point. He particularly underlines that we can find many examples for discrete manifolds, whereas continuous manifolds are rare. Yet, the latter shapes the field of higher mathematics in which Riemann's notion of manifold serves a fundamental role.

Here Herbart offers a form of intelligible geometry that is compatible with Riemann's approach to investigating the foundations of geometry. Both philosophers think that in order to investigate the foundations of geometry we have to avoid considering space as given; instead they claim that we have to construct geometry starting from basic concepts. While Herbart claimed that this construction was possible from any continuum, Riemann adopts a scientist's point of view and says that we can view any space as an n-fold extended manifold, where n is the number of independent directions in which we can travel. Riemann's ontology concerning mathematics can best be understood in connection with Herbart's view of mathematics. Herbart regarded mathematics as part of philosophy because he thought that, like philosophy, mathematics turns its *concepts* to *its subjects*; this is a process that goes far beyond the manipulation of formulas (Scholz 1982, 425). Riemann uses the term *speculation* in trying to solve problems. Philosophy makes use of speculation, and its subjects are concepts. In the context of formation, development, and extension of scientific concepts Riemann sees the position of mathematics similarly to the role Herbart ascribes to philosophy. Herbart thought that the sciences developed their central concepts with respect to their contexts; however, philosophical studies of the sciences require more; they must form unifying concepts that transcend this or that specific context (Scholz 1982, 424). These ideas seem to influence Riemann's ideas about the methodology of mathematics. Riemann's studies in different fields of mathematics (complex function theory, geometry, and integration) show that he wanted to develop and use his geometric ideas on n-dimensional manifolds. Diversity in geometric thought could be kept together or, to put it in more philosophical language, it could be represented as 'a unity in diversity'. Riemann does this with the concept of manifold, for it could admit different enrichments in order to show the *possibilities* and *conceptual freedom* of geometric thought (Scholz 1992, 4). Riemann understands science as 'the attempt to perceive nature through accurate concepts'. Riemann's understanding of *concept* must be interpreted in accordance with his main aim, which was to perceive nature that is dynamic in character. The only way to grasp nature and its changing character is to study it, that is, by adjusting and modifying our concepts with respect to nature—which means that our concepts cannot be given, fixed, or necessary. In this sense the concept of space cannot be an exception; rather it must be an instance of 'multiply extended magnitude' that is capable of change and variation. For Riemann, this means abandoning *a priorism* and emphasizing the role of *hypotheses.*

Herbart's influence on Riemann is seen in the epistemology and conceptual methodology of mathematics. Thus, I think that the crucial point Riemann took from Herbart when liberating geometrical thought is the idea that we do not necessarily identify physical space with the space of the senses. Riemann aimed at founding geometry anew on our perception and on the construction of space. Hence, the first part of the *Habilitationsvortrag* lecture of 1854 reflects the philosophical investigations that influenced Riemann. The concept of manifold is philosophical since it is a concept that enabled Riemann to show the possibility of other geometries and examine the *necessity* and *a priority* of Euclidean geometry.

2.3. Gauss on the Nature of the Space

Gauss was opposed to the Kantian conception of space and geometry. For Gauss, space must have a real meaning. He made this point in a letter in response to Bolyai:

> Precisely in the impossibility of deciding a priori between Σ (the Euclidean system) and S (the system of the science of space) that we find the clearest demonstration that Kant was wrong to state that space is only a form of our intuition. Another and just as strong reason I have had occasion to point out in a short note in the *Göttingischen gelehrten Anzeigen 1831.* (Gauss in Bottazzini 1994, 23)

'Strong reason', Gauss points out, refers to Kant's *incongruent counterparts*; an argument Kant thinks shows *a priori* nature of space.[8] According to Gauss:

[8] In his transition to his *Critical* period, advancing this argument Kant claims that since our right and left hands have Leibnizian internal spatial relations we can think of them as equal. However, since we cannot superimpose our one hand upon the other they are incongruent. Then there arises a difference between them concerning not to the spatial relations among their parts but space itself. Kant seems

> This difference between right and left is in itself completely determined soon as a random front and back have been fixed on a plane and an above and below in relation to the surfaces of the plane; only if we change our intuition of this difference can we communicate it by indicating really existing material objects. (Gauss in Bottazzini 1994, 23)

Although Gauss agrees with the premises of Kant's argument, he believes that, contrary to Kant, they prove that 'space is not an a priori form of intuition'. That is, if anything, these premises prove that space must have a real physical meaning, i.e., 'space, regardless of our capacity of intuition, must have a real meaning' (Gauss in Bottazzini 1994, 23)

To search for evidence that the geometry of space is non-Euclidean, in the early 1820s Gauss measured the angles of a large triangle formed by light rays joining three peaks.[9] In emphasizing the importance of empirical investigation in *Habilitationsvortrag*, Riemann clearly reflects this Gaussian heritage. In addition, Riemann's strong interest in physics is clear from the fact that he was the physicist Weber's assistant for eighteen months. Although Riemann was under the influence of Gauss and therefore Gauss' empirical approach to geometry, it is nevertheless hard to call him a thoroughgoing empiricist. Before defining any metrical relations, the possibility of different geometries had to be investigated on the basis of the concept of manifold. In doing so, axioms of Euclidean Geometry are not only the 'most important', they are also empirically contingent rather than logically necessary, so that 'one may therefore inquire into their probability'. This shows that Riemann's main aim was not to reinterpret or to modify previously given geometric knowledge, and nor was it to examine the classical questions; rather, his main aim was to expand the domain of geometry—by which he would open new vistas for physical thought (Ferreiros 2007, 60-

---

to argue in favor of Newton's absolute as opposed to Leibniz's relational space. However, there is no agreement concerning whether or not this is a valid interpretation of the purpose of argument. See, for example DiSalle (2006, 62-63).

9 This issue sparks significant controversy among historians of science, as evident in the debate between Arthur Miller's "The Myth of Gauss' Experiment on the Euclidean Nature of Physical Space" (1972) and Erhard Scholz's (2006) response to it.

61). That is why Riemann wants to speak of *hypotheses*, rather than *axioms*, in his lecture.

2.4. Mathematics in Habilatitionsvortrag

In the mathematical part of his *Habilatitionsvortrag*, Riemann follows Gauss' most fundamental steps, by extending Gaussian concepts and results for surfaces to n-dimensional manifolds, such as the measure of curvature and some properties of geodesic lines. Like Gauss, Riemann's approach is metric; the concept of distance plays a fundamental role both in the theory of curved surfaces and in Riemannian manifolds; in addition, the essential properties of manifolds are expressed by means of the linear element. Gauss' treatment of curved surfaces is of special importance in Riemann's *Habilitationsvortrag.*[10] Specifically, *Disquisitiones Generales Circa Superficies Curvas* of 1828, in which Gauss introduces his *Theorema Egregium* ('Remarkable Theorem'), includes all the results and concepts that are later developed and extended by Riemann. In his studies on surfaces, Gauss had already reached a formula that is then developed and extended by Riemann. Riemann's starting point was the equation $ds^2 = Edu^2 + 2Fdudv + Gdv^2$ where a point determined by coordinates u and v on a surface in Euclidean space and E, F, and G are functions of the variables u and v. Based on this formula, we say that if we know the curvature then all the measure relations can be determined universally, that is, by means of Gaussian curvature we reach what can be called 'invariant structure'. Metric coefficients' behaviour on a surface contains all the information about the geometry of the surface. Without reference to any space outside the surface it is possible to know the measure of the curvature at the point determined by u and v, as a function of E, F, G, and their differentials. To put it another way: Gauss

[10] Riemann, potentially influenced by his mentor Gauss, pioneered the study of curved surfaces coordinated by parameters and introduced metrical-differential concepts. However, it is important to note that while Gauss played a significant role in advancing intrinsic geometry, it is worth noting a pre-existing tradition of mathematical endeavours focused on metrics, curved surfaces, and differential concepts. For instance, Leonhard Euler's studies could be considered within this tradition. See Papadopoulos (2017).

showed that we can do geometry on a surface (two-dimensional) independently of surrounding Euclidean (three-dimensional) space.

In *Habilitationsvortrag*, Riemann generalizes the Gaussian theory of curved spaces to n-dimensions. Such manifolds are characterized by the fact that each point within them can be uniquely specified by n real numbers. The introduction of the concept of distance into a manifold follows the Gaussian model. Analogously to the two-dimensional case, infinitesimal distances are expressed by processing differentials given in terms of some internal coordinate system, u, with the help of the metric tensor g_{ij}. Thus, Riemann arrives at a formula that is identical to the Gaussian expression for the surfaces:

$$ds^2 = \sum_{ij} g_{ij} dx^i dx^j \quad (1)$$

where g_{ij} are functions of coordinates, and $x^1, \ldots, x^n$ are coordinates on manifold. This quadratic form satisfies the following conditions:

a) the metric is symmetric, ($g_{ij} = g_{ji}$)
b) Positive definite matrix for all $1 \leq i, j \leq n$, which are the basic conditions to measure the distance in Euclidean space.

Although Gauss' studies of surfaces led to the discovery of the intrinsic aspects of space, he had nevertheless worked with Euclidean space—in fact, his surfaces were embedded in Euclidean space. However, Riemann—by attaching to each point of the manifold a Euclidean tangent space—in contradistinction to Gauss did not make use of the notion of an embedding space of higher dimensionality. Equation *(1)* brings about important results. A manifold thus allows for a distinction between neighbouring points—or events, in case of a space-time manifold—and distant points or events. The first point comes in relation to methodology of work. It brings to geometry the approach of theoretical physics: since differential expression *(1)* allows point by point analysis, it is useful for constructing basic laws that govern complex processes occurring within infinitely small elements of space or time. Second, *(1)* shows the possibility of different geometric systems, each of which depends on the chosen metrical system employed on the same manifold. In these different geometric systems of n-dimensions, geometric relations are no different from classical geometry. Although *(1)* is very general and also includes manifolds whose curvature is variable from one point to another, since Riemann finds these manifolds more interesting from a theoretical point of view, he chooses to consider a manifold of constant curvature—any portion of which can be continuously superposed anywhere on the manifold.

2.5. Physics in Habilatitionsvortrag

The last part of the *Habilitationsvortrag*, 'Applications to Space', develops an analysis of Euclidean space that Riemann characterizes as a three-dimensional flat space with curvature equal to zero. In this section, Riemann considers the necessary and sufficient conditions for determining metric relations in space. Riemann's three conditions are not independent but offer alternative characterizations of Euclidean Space:

1) The angles of triangles define it (at every point and in all planar directions).
2) It can be defined by the concept of "free mobility" of a rigid body at each point and in all directions, where assuming zero curvature at one point suffices.
3) It can also be characterized by the integrability of direction and length, making it not just a Riemannian manifold but a global vector space, where vectors' directions and lengths have meaningful distinctions.

The crucial aspect is not merely the existence of lines or bodies but their ability to be rigidly moved or to have their measures (length, area, volume) defined independently of their position and orientation in the manifold.

'Applications to Space' shows that for Riemann geometrical structures have fundamental physical significance in that they allow us to perceive nature in a more intelligible manner. Properly chosen, an infinite variety of geometric systems can be defined, which can function as tools for studying natural phenomena through their geometric representations. Thus, Riemann saw a strong relationship between geometry and the image of the physical universe. The physical importance of the concept of manifold is proven by a ground-breaking General Theory of Relativity. In General Relativity, space is conceived as a four-dimensional differentiable spacetime manifold (simply our ‘world’), in which metric is determined by the matter (Boi 1992, 198). As a result, Einstein’s principle of equivalence unifies metric and gravitation. We see the line element of a Riemannian manifold again:

$$ds^2 = \sum_{i,k=1}^{n} g_{ik} dx^i dx^k \quad (g_{ik} = g_{ki})$$

In General Relativity the function g_{ik} denotes the gravitational field. However, although the mathematics of Gauss and Riemann paved the way for Einstein’s Theory of General Relativity, it would be an overstatement to say that Riemann had foreseen the meaning, in physical terms, of his generalization of geometry. Riemann did not foresee what Einstein later accomplished. What he saw was not the emergence of a four-dimensional spacetime, but rather an understanding of the usual three dimensions of physical space as a particular case of n-dimensional space (Ferreiros 2004, 1). In *Habilitationsvortrag*[7] he says the main applications of his ideas would not be found in the large, but rather in the extremely small, since for him most of the physical phenomena on the microscopic level could not be explained by Euclidean light rays and rigid body, as at this level bodies would no longer exist independently of place, and because curvature of space would no longer be constant.[11]

2.6. Riemann’s Philosophy of Geometry

Interestingly, in *Habilatititionsvortrag* Riemann does not use the term ‘non-Euclidean geometry’, and he does not refer to the studies of János Bolyai or Nikolai I. Lobatchevsky; nor does he try to compare his views to Kant’s philosophy of space.[12] Although Riemann probably knew the studies

---

[11] For a discussion of older literature on Riemann’s *Habilitationsvortrag*, see Nowak (1989).

[12] Laugwitz connects Riemann’s avoidance of a direct attack on Kant with the presence of Rudolph Hermann Lotze among the audience of the *Habilitationsvortrag.* Lotze, as a follower of Kantian tradition, opposed non-Euclidean geometry—arguing that it is nonsense. See Laugwitz (1999, 222). Botazzini (1994, 25) claims that the *Habilitationsvortrag* was delivered in order for Riemann to qualify as a *Privatdozen*t, hence in such a delicate examination it would be better for Riemann to not enter into a discussion on such a controversial subject.

of Bolyai or Lobatchevsky (and of course Gauss), he cautiously refrains from discussing Bolyai's and Lobatchevsky's approaches to constructing non-Euclidean geometry, which involved an axiomatic method that negated Euclid's fifth postulate.[13] Yet Riemann's geometry requires us to at least re-evaluate our philosophical theories of geometry based on Euclidean geometry. For Kant, space and time are not concepts, they are pure forms of intuition (*Anschaauung*), and space is uniquely determined by three-dimensional Euclidean geometry and its propositions. Let us recall Kant's core arguments in the *Critique of Pure Reason*, about space. Space is the source of all synthetic a priori propositions of geometry. It is empirically real, but transcendentally ideal. It is a necessary condition of all objective experience, but it has no existence outside of our experience. All experience of objects in spatial relationships presupposes a space in which they are ordered. Space is an a priori intuition. We cannot represent to ourselves the absence of space. Space is not a concept of the relation of things. (a) There is only one space; and (b) the parts cannot precede this whole since they must exist within it. In addition, the synthetic a priori propositions of geometry are only possible if space is an a priori intuition. Space is an infinite given magnitude. No concept of relations can give rise to infinitude and no concept can contain an infinite number of representations within it.

It seems that Riemann could not agree with any of these propositions. Riemann's point was, as the structure of the *Habilitationsvortrag* clearly shows, that instead of postulating the axioms of Euclidean geometry, we should consider the conjunction of those axioms with a physical interpretation, and ask whether they were in point of fact really true. In Riemann's view, space has a physical reality. It is something given in experience together with metric determination. For pragmatic reasons he creates different spaces, and the question of what kind of geometry is true of space is a question of empirical determination, and is thus *a posteriori*.

---

[13] Although he does not mention these names he probably knew their works. One of the works of the Bolyai was presented in Crelle's journal *The Journal für die reine und angewandte Mathematik* (Journal for pure and applied mathematics) in 1837 (Bottazzini & Tazzioli 1995, 27). In addition, it is highly probable that he could have gained knowledge about the geometries of Bolyai and Lobatchevsky through Gauss (Laugwitz 1999, 224).

In my view, the *Habilitationsvortrag* is a perfect example of the interplay between philosophy and mathematics. Herbart's constructive approach to space inspired Riemann to create a fruitful combination of higher-dimensional geometry and Gauss' differential geometry (Banks, 2013). Riemann also followed Herbart and Gauss in rejecting Kant's view of space as an a priori form of intuition. Riemann regards space as a concept with meaning for the physical realm and as capable of change and variation. In accordance with Herbartian psychological theory, Riemann adopts a materialist criterion of truth and wants to answer the question: 'When is our conception of the world true?' with 'When the coherence of our concepts corresponds to the coherence among things', and when the 'connection of things' is deduced from 'connections of phenomena' (Riemann as quoted in Ehm 2010, 145). Throughout the *Habilitationsvortrag*, we see that the investigative hypothesis lying at the basis of geometry, and which was Riemann's main concern, was infinitesimals. Developing this approach enabled Riemann to investigate the links between different laws of nature—knowledge of which is based on the exactness of our description of phenomena in infinitesimal regions. Gaining knowledge of the external world from the behaviour of infinitesimal parts constitutes the backbone of Riemann's research program. In fact, *Habilitationsvortrag* of is a summary of Riemann's metric approach, which aimed to find the concept of an n-dimensional manifold equipped with the notion of distance between infinitely close points.

2.7. Kant and Riemann on Pure Intuition

Despite all the reasons for thinking that Riemann's philosophy of geometry and Kant's spatial intuition do not go together, I would like to argue that they are not necessarily inconsistent. In *Habilatitionsvortrag*, Riemann maintains that the main principles that lie at the foundation of geometry are hypotheses, and that their value is determined within 'the bounds of observation'. Here I want to underline the phrase 'the bounds of observation'. Riemann stresses that there exists some form of limit, which may well be the same as the perceptual capacity given in Kant's spatial intuition. For Kant, space and time, which are *the forms of pure intuition*, are not *concepts*. That is, one should be able to use concepts as *predicates* of subjects; but with 'space' and 'time', such predication is not possible. We can talk of the spatiality and the temporality of a thing, but while doing so, what we talk about are not space and time, but the parts that are derived from space and time. We don't conceive the external world as islands of space-time; instead we conceive it within the integrity of space-time—if we think of space and time as *complete* and *unique* then the suggestion that space and time are concepts can be ruled out. In addition, with the help of the idea of 'incongruent counterparts', we can understand why 'space' and 'time' are the pure forms of intuition but not concepts. Since we cannot define our right and left hands, we cannot name them on a conceptual level. Thus, for Kant space and time function as the conditions of all possible experience. A close reading of Kant suggests that he did not say space had to have the properties described in Euclidean geometry; rather, and at most, that we necessarily perceive space *as if it were Euclidean*. Kant's point was that as humans, or perhaps as living beings, we perceive space in some geometric system, simply by virtue of being human. The Euclidean system introduces geometrical constraints. It is true that Riemann introduces his concept of manifold in a rather quasi-philosophical way. However, according to Michael Spivak, Riemann was clear that manifolds are, locally, similar to n-dimensional Euclidean space:

> However, it is quite obvious that the notion was thoroughly clear in his own mind and that he recognized that manifolds were characterized by the fact that they are locally like n-dimensional Euclidean space. (1975, 155)

Riemann thought that geometry must start from infinitesimals. The metric given by the standard Euclidean distance $ds^2 = \sum_{ij} \delta_{ij} dx^i dx^j$ in n-dimensional Euclidean space $\mathbb{E}^n$ is the same distance relation as $\mathbb{R}^n$.[14] In Riemann's characterization of n-dimensional curvature a region of manifolds

[14] Here it is important to note that we do not say that the Pythagorean Theorem holds in every Riemannian manifold; rather, what we try to say is that by means of the notion of manifold we can transport some known theorems of Euclidean operations to n-dimensions. For example, the Pythagorean Theorem is valid in both $\mathbb{R}^2$ and $\mathbb{R}^n$. In these different geometric systems of n-dimensions, relations are no different from those in classical geometry. To give an example, in the classical Euclidean system in two dimensions, we employ the Pythagorean Theorem $a^2 + b^2 = c^2$, while

counts as flat if the distance between any pair of points in it satisfies Euclidean metric. Kant's claim about space being an *a priori* form of pure intuition, and Riemann's point about intuitive space in this sense do not necessarily rule each other out. Here it is crucial to distinguish between the foundational topological structure of a Euclidean space and the space itself, which, according to definition, constitutes a metrical space. This differentiation holds significant philosophical importance in grasping the core structure of Riemann's dissertation. He establishes the metric as a differential quadratic form based on certain *a priori* assumptions at the outset of the second part of his dissertation. It's essential to avoid conflating these with the empirical determinations of metric coefficients in the third part. Specifically, empirical evidence would not be capable of discerning between a quadratic form and a more intricate metrical function, such as one derived from a fourth-degree homogeneous form. I think that for Riemann topological structure is unique and necessary but metrical structure is subject to empirical investigation. Riemann aimed to establish the foundational essence of the metric as a differential quadratic form, particularly evident in the early part of his *Habilatitionsvortrag*'s second section. Specifically, concerning Riemannian manifolds, the ability to adjust metric coefficients for various physical applications does not imply that the fundamental concept itself—the notion of a differential quadratic form—is subject to fluctuation. Indeed, Hermann Weyl clearly distinguishes within a Riemannian manifold between two aspects: 1) the fundamental concept of a differential quadratic form, which Riemann establishes *a priori* in the construction outlined in the second part of his dissertation—Weyl termed this "the essence of space/the metric" (1923, 102-103); and 2) the contingent variation of the metric coefficients, which is subject to change and can be linked to empirical investigations. Similarly, Ernst Cassirer (1923) recognized the clear division between the concept of the Riemannian metric itself and the variation of its coefficients. He proposed that the overarching idea of a Riemannian manifold, when considered in its entirety with variable coefficients, could be regarded as the fixed *a priori* space essential for relativistic physics. Hence, Riemann's line of reasoning both adheres to and extends beyond Kant's. For all these reasons, we can

in three dimensions it takes the form of $a^2 + b^2 + c^2 = d^2$, and in n-dimensional case it will become $x_1^2 + x_2^2 + \cdots + x_n^2 = z^2$.

perhaps call Riemann's philosophy of geometry a neo-Kantian philosophy of geometry.

As Luciano Boi puts it, Kant suggests that one role of the intuition of space in external sensibility is to lend a rational and structural coherence to various empirical phenomena. However, this coherence appears to rely less on the internal structure of the phenomena themselves and more on a faculty inherent to subjectivity. In essence, intuition functions as a frame of reference or an organizing principle (2019, 3). Indeed, Boi quotes Riemann in order to show Riemann agrees on and even more precise concerning this function of spatial intuition:

> The hypothesis that space is an infinite and three-dimensional manifold is a hypothesis which applies to our whole perception of the external world, and which allows us, at every instant, to complete the realm of our actual perceptions and construct the possible places of objects; in fact, this hypothesis is constantly confirmed in all of these applications. (...) But the infinitude of space is by no means a necessary consequence of what precedes. (Riemann, 1990)

According to Kant, in order to represent to oneself various kinds of spaces, all of which are logically possible, one needs first to possess the concept of space. Riemann's concept of manifold can actually be thought as this concept of space that Kant thought was necessary for representing various kinds of spaces to ourselves. Riemannian manifolds can represent non-Euclidean spaces, each of which is dependent on the chosen metrical system employed on the same manifold. Thus, on this view of Riemann's philosophy of geometry, spatial intuition is not being dismissed. What is more, Riemann—rather than being in opposition to Kant—shows that there are valuable conceptual resources to be found when applying geometry to physics in the very large and the very small.

In *Habilitationsvortrag*, Riemann also makes a distinction between 'unlimitedness' and 'infiniteness'. This distinction can be understood in light of a distinction between the qualitative features of space, i.e; 'the extent relations', and features relating to distance, i.e., 'measure relations'. In more modern terms, 'relations of extension' correspond to 'topological relations', while 'measure relations' correspond to 'metrical relations'. We can see this distinction when considering the surface of a sphere: it is not infinite in extent but unbounded. For Riemann, properties such as unboundedness and three dimensionality of space are known with an empirical certainty greater than that of any experience of the external world:

> That space is an unlimited, triply extended manifold is an assumption which is employed for every apprehension of the external world; by it at every moment the domain of real perceptions is supplemented and the possible locations of an object that is sought for are constructed, and in these applications the assumption is continually being verified. (1929/1959, 423)

Riemann's stress on unboundedness is followed by a question about whether our certainty about unboundedness is compatible with our certainty about the infinitude of space. I think it is not inappropriate to claim that, for Riemann, when we say that 'space is a three-dimensional manifold', has the same empirical certainty as the statement 'it is unbounded'. The above quote shows that, for Riemann, our perception of the external world is limited to three-dimensional Euclidean geometry—but he makes no reference to Kantian spatial intuition in this context. Kant argues for the intuitive nature of space at B40 in *Critique of Pure Reason* by appealing to its unboundedness—where the unboundedness of space is supposed to be guaranteed by our prior recognition. The idea that on a single topology many metric relations are possible can be used to interpret Kant's understanding of space. In the *metaphysical exposition* there are four main propositions about space: Space is not an empirical concept which has been derived from outer experience. (1781/2007, B38, 68)

The basic idea here is that if the representation of space is presupposed then relational aspect of things is possible: 'Space is a necessary a priori representation which underlies all other intuitions'. (1781/2007, A24/B39, 68) Space is the sole requirement of the possibility of external appearances; therefore, it must be an *a priori intuition*: 'Space is not a discursive or, as we say, general concept of relations of things in general, but a pure intuition'. (1781/2007, A25, 69) Here Kant argues that we can represent to ourselves only one space, and we can only consider parts of this unique space. Parts cannot precede this whole space; therefore, they can only exist *in* this space; and therefore, space is necessarily one and is an *a priori intuition:*

'Space is represented as an infinite given magnitude'. (1781/2007, B40, 69) Kant's idea here is that a concept can have infinitely many different representatives as instances of it, but the concept itself cannot be represented in infinitely many different ways. Every concept contains infinitely many representations under itself, but not within itself. Therefore, space can only be thought of in this latter way. As such, the original representation of space is not a concept, but an *a priori* intuition. In this metaphysical exposition, I do not think that Kant is giving a topology of space any different to the concepts of being unbounded and of continuous intuition. Torretti (1984, 33) suggests that:

> Since Kant conceived the 'manifold of a priori intuition' called space, not as a mere point-set, but as a (presumably three-dimensional) continuum, we must suppose that he would expected 'the mere form of intuition' to constrain the understanding to bestow a definite topological structure on the object of geometry. But, apart from this, the understanding may freely determine it, subject to no other laws than its own. Since the propositions of classical geometry are not logically necessary, nothing can prevent the understanding from developing a variety of alternative geometries (compatible with the prescribed topology), and using them in physics.

Hence, based on the idea that Kant does not give a unique determination of space, it is possible to argue that any possible space would have a geometrical structure that is not graspable by human understanding. Yet topological properties, such as continuity, three dimensionality, and unboundedness count as constraints directly imposed by the mere form of intuition. Riemannian manifolds are compatible with constraints imposed by Kantian spatial intuition in a topological sense.

Kant's pure intuition (*reine Anschauung*) is one with which we represent ourselves in physical space. Since we can think of empty space, but not the absence of the space, the concept is *a priori*. Second, for Kant space as pure intuition is the same as the physical space of ordinary experience—that is, empirically real and transcendentally ideal. According to some commentators (Wiredu 1970; Friedman 1992 and 1999), it is also possible to talk about logically possible space, for Kant. So, what is logically possible space for Kant? Kant distinguishes between the logical possibility of a concept and the objective reality of a concept:

> Thus there is no contradiction in the concept of a figure which is enclosed within two straight lines since the concepts of two straight lines and their coming together contain no negation of a figure. The impossibility arises not from the concept itself, but in connection with its construction in space, that is, from the conditions of space and its determination.[15] (1781/2007, A221/B268, 240)

Kant argues that such a concept is not self-contradictory or logically possible and is in fact objectively real. He defines the possible as that which is objectively real; and he refers these concepts as 'fictitious', denying that they tell us anything about space. Thus, Kant equates intuited space with physical space, and for this reason he thinks that logically possible spaces are not really informative about space. Riemann argues along the same lines. After pointing out that the simplest case for space is determined by $ds^2 = \sum_{ij} g_{ij} dx^i dx^j$, he says that:

> The next case in order of the simplicity would probably contain the manifolds in which the line-element can be expressed by the fourth root of a differential expression of the fourth degree. Investigation of this more general class indeed would require no essentially different principles, but would consume considerable time and throw relatively little new light upon the theory of space, particularly since the results cannot be expressed geometrically. (1929/1959, 417)

[15] This quotation combines Kant's logical criterion of possibility with Friedman's assertion that Kant's notion of real possibility can be replaced with our notion of physical possibility. Friedman's point is that in Kant's distinction between conditions of thought and conditions of cognition, the former does not correspond to our notion of logical possibility—rather, logical possibility as given by the conditions of thought plus intuition corresponds to pure mathematics. On the other hand, real possibility as given by the conditions of thought plus empirical intuition corresponds to the '(pure part of) mathematical physics' (Friedman 1992, 94).

Riemann's strategy is to identify intuited space as just one logically possible space, and not necessarily as a true description of physical space (Nowak 1989, 20). What's new here is an altered definition of space as manifold, which eliminated the necessity for a definition of three-dimensional Euclidean space and, by implication, the necessity for propositions using concepts in Euclidean geometry (propositions that were formed out of the concepts). In doing so, Riemann shows the possibility of getting rid of the 'necessity' of concepts in a system of geometry (concepts being necessary for a system). For Riemann, the concept of manifold (n-dimensional topological space) is the most general structure common to this infinite multiplicity of spaces. In this sense, he is able to identify logically possible geometries, just as Kant had suggested. As such, we can say that for Riemann this structure represents the general condition for the perception of the matters of fact and, therefore, that it is the *a priori* form of spatial intuition. Thus, for Riemann there is an *a priori* intuition of space, which is not metrical but topological. We can therefore say that what Riemann's philosophy of geometry denies is Kant's claim about the equality of intuited space with physical space; and not Kant's point about the *a priori* necessity of a general concept of space for any theory of geometry.

3. Conclusion

Riemann's ideas seem to have strong philosophical implications, yet I don't think that they were developed against Kant's philosophy of geometry. Rather, these concerns guided him to find a satisfactory basis for studying nature from more general point of view. Revolutionizing mathematics and physics was not what Riemann intended. He wanted to deal with a problem that had been around for a while—namely, is there something besides Euclid? How can we be sure about Euclid's axioms? Riemann suggests a procedure for being sure; he says that we can view any space as an n-fold extended manifold, where n is the number of arbitrary directions in which we can go. Thus, the problem for Riemann is different; that is, Riemann had no real interest in the problem of the foundations of geometry as such, such that the problem of parallels belongs to the foundations of elementary geometry—yet as his 1851 dissertation shows, he wanted to develop and use geometric ideas on n-dimensional manifolds as an aid to mathematics and physics. Riemann was fundamentally interested, not in synthetic *a priori* propositions, but in the geometry of physical space. In this sense, we can only say that Riemann had no particular commitment to the truth of Euclidean geometry; yet this does not mean that he wanted to cast doubt upon the Euclidean axioms. Rather, he had the goal of developing geometry so that it would become accessible to science and empirical verification.

References

Banks, Eric C. 2005. "Kant Herbart Riemann. " *Kant Studies*, 96(2): 208–34. 10.1515/kant.2005.96.2.208

Banks, Eric C. 2013. "Extension and Measurement: A Constructivist Program from Leibniz to Grassman." *Studies in History and Philosophy of Science*, 44(1): 20-31. https://doi.org/10.1016/j.shpsa.2012.07.011

Barsan, Victor, Merticariu, Andrei. 2016. "Goethe's Theory of Colors between the Ancient Philosophy, Middle Ages Occultism and Modern Science." *Cogent Arts & Humanities*, 3(1): 1–29. https://doi.org/10.1080/23311983.2016.1145569

Boi, Luciano. 1992. "The 'Revolution' in the Geometrical Vision of Space in the Nineteenth Century, and the Hermeneutical Epistemology of Mathematics." In *Revolutions in Mathematic*, edited by Donald Gillies, 183–208. Oxford University Press. https://doi.org/10.1093/oso/9780198539407.003.0011

Boi, Luciano. 2019. "The Role of Intuition and Formal Thinking in Kant, Riemann, Husserl, Poincare, Weyl, and in Current Mathematics and Physics." *Kairos. Journal of Philosophy & Science*, 22 (1) 1–53. https://doi.org/10.2478/kjps-2019-0007

Bottazzini, Umberto.1994. "Geometry and Metaphysics of Space in Gauss and Riemann". In *Romanticism in Science,* edited by S. Poggi, M. Rossi, 15-29. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-017-2921-5_2

Bottazzini, Umberto. 1995. "Naturephilosophie and Its Role in Riemann's mathematics." *Revue d'histoire des mathématiques* 1, 3–38. https://doi.org/10.24033/rhm.1

Cassirer, Ernst. 1923. *Substance and Function and Einstein's Theory of Relativity*, translated by Marie Collins Swabey and William Curtis Swabey. Chicago: Open Court. https://doi.org/10.2307/2179226

Laugwitz, Detlef. 1999. *Turning Points in the Conception of Mathematics: Bernhard Riemann 1826–1866.* Boston: Birkhauser. https://doi.org/10.1007/978-0-8176-4777-3_5

DiSalle, Roberto. 2006. *Understanding Space–Time*, Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511487361

Ehm, Werner. 2010. "Broad Views of the Philosophy of Nature: Riemann, Herbart and the Matter of the Mind." *Philosophical Psychology*, 23(2): 141–62. https://doi.org/10.1080/09515081003690400

Friedman, Michael. 1985. "Kant's Theory of Geometry." *Philosophical Review*, 94(4): 455-506. https://doi.org/10.2307/2185244

Friedman, Michael.1992. *Kant and Exact Science*. Cambridge, Massachusetts: Harvard University Press. https://doi.org/10.2307/2185822

Friedman, Michael.1999. *Reconsidering Logical Positivism*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139173193

Ferreiros, Jose. 2004. "The Magic Triangle: Mathematics, Physics and Philosophy in Riemann's Geometrical Work." Retrieved from http://www.docdatabase.net/more-the-magic-triangle-mathematics-physics-and--961738.html (accessed December 10, 2020).

Ferreiros, Jose. 2007. *Labyrinth of Thought: A History of Set Theory and its Role in Modern Mathematics.* 2nd ed. Basel and Boston: Birkhauser. https://doi.org/10.1007/978-3-7643-8350-3

Gray, Jeremy. 1989. *Ideas of Space*, Oxford: Clarendon Press.

Kant, Immanuel. 1985. *Prolegomena*, translated by J.W. Ellington. Indianapolis: Hackett.

Kant, Immanuel. 1986. *Logic*, translated by R. Hartmann and W. Schwarez. New York: Dover.

Kant, Immanuel. 2004 [1786]. *Metaphysical Foundations of Natural Science*, translated and edited by M. Friedman. Cambridge: Cambridge University Press.

Kant, Immanuel. 2007 [1781]. *Critique of Pure Reason*, translated by N.K. Smith. London: Palgrave Macmillan.

Lenoir, Timothy. 2006. "Operationalizing Kant: Manifolds, Models, and Mathematics in Helmholtz's Theories of Perception." In *The Kantian Legacy in Nineteenth-Century Science*, edited by M. Friedman and A. Nordmann, 141–210. Cambridge, Massachusetts: MIT Press.

Miller, Arthur. I. 1972. "The Myth of Gauss' Experiment on the Euclidean Nature of Physical Space." *Isis*, 63(3): 345–48. https://www.jstor.org/stable/229274

Morales, Alberto. 2019. "From Grassmann, Riemann to Husserl: A Brief History of Concept of Manifold." *Meta: Research in Hermeneutics, Phenomenology, and Practical Philosophy*, XI(2): 473–500.

Nowak, Gregory. 1989. "Riemann's Habilitationsvortrag and the Synthetic *a priori* Status of Geometry." In *The History of Modern Mathematics*, vol. 1, edited by David E., 17–48. Boston: Rowe-Academic Press

Plotnitsky, Arkady. 2009. "Bernhard Riemann's Conceptual Mathematics and the Idea of Space." *Configurations*, 17(1): 105–30. 10.1353/con.0.0069

Riemann, G. Bernhard. 1929/1959. "On the Hypotheses Which Lie at The Foundations of Geometry." In *A Source Book in Mathematics*, vol. 2, edited by D.E. Smith, 411-25. New York: Dover.

Riemann, G. Bernhard. 1990. "On the Hypothesis Which Lie at the Basis of Geometry". In *Gesammelte Mathematische Werke/Collected Papers*, edited by R. Narasimhan. Berlin: Springer.

Russell, Bertrand. 1956. *An Essay on the Foundations of Geometry.* New York: Dover.

Scholz, Erhard. 1982. "Herbart's Influence on Bernhard Riemann." *Historia Mathematica*, 9: 413–40. https://doi.org/10.1016/0315-0860(82)90106-9.

Scholz, Erhard. 1992. "Riemann's New Vision of a New Approach to Geometry." In *1830–1930: A Century of Geometry*, edited by D.F.L. Boi, 22–34. Berlin: Springer.

Scholz, Erhard. 2006. "Another Look at Miller's Myth". Preprint Wuppertal.

Spivak, Michael. 1975. *A Comprehensive Introduction to Differential Geometry*, vol 2, 2nd ed. Houston: Publish or Perish, Inc.

Tazzioli, Rossana. 2003. "Towards a History of the Geometric Foundations of Mathematics Late XIXth Century." *Revue de Synthèse*, 124(1): 11-41.

Torretti, Roberto. 1978. *Philosophy of Geometry from Riemann to Poincare*. Dordrecht, Boston and Lancaster: D. Reidel.

Ohshika, Ken'ichi. 2017. "The Origin of the Notion of Manifold: From Riemann's Habilitationsvortrag Onward." In *From Riemann to Differential Geometry and Relativity*, edited by Ji L., Papadopoulos A., Yamada S., 295-309. Cham: Springer.

Papadopoulos, Athanese. 2017. "Looking Backward: From Euler to Riemann". In *From* Riemann to Differential Geometry and Relativity, eds. Ji L., Papadopoulos A., Yamada S.,1–94. Springer, Cham.

Plotnitsky, Arkady. 2017. "Comprehending the Connection of Things: Bernhard Riemann and the Architecture of Mathematical Concepts". *From Riemann to Differential Geometry and Relativity*, edited by Lizhen Ji, Athanase Papadopoulos, and Sumio Yamada, 329-363. Cham: Springer.

Weyl, Hermann. 1923. *Raum, Zeit, Materie: Vorlesungen über allgemeine Relativitätstheorie*, 5th edition. Berlin: Springer.

Wiredu, Kwasi. 1970. "Kant's Synthetic a Priori in Geometry and the Rise of Non-Euclidean Geometries." *Kant-Studien*, 61(1): 5–27. https://doi.org/10.1515/kant.1970.61.1-4.5

Organon F 31 (2) 2024: 141–182
https://doi.org/10.31577/orgf.2024.31203
ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

How to Defend the Law of Non-Contradiction without Incurring the Dialetheist's Charge of (Viciously) Begging the Question

Marco Simionato*

Received: 10 August 2023 / Revised: 11 March 2024 / Accepted: 15 April 2024

Abstract: According to some critics, Aristotle's *elenctic* defence (*elenchos*, *elenchus*) of the Law of Non-Contradiction (*Metaphysics* IV) would be ineffective because it viciously begs the question. After briefly recalling the elenctic *refutation* of the denier of the Law of Non-Contradiction, I will first focus on Filippo Costantini's objection to the elenchus, which, in turn, is based on the dialetheic account of negation developed by Graham Priest. Then, I will argue that there is at least one reading of the elenchus that might not be viciously question-begging. In doing so, I will leverage, reinterpret and adjust the distinction between two senses of epistemic dependence, offered by Noah Lemos and originally based on some thoughts about George Edward Moore's 'proof of an external world.' The key point of my counter-objection to recover the elenchus is to use the distinction between a necessary-condition relation between propositions (p only if q) and a grounding relation between facts (the fact that an epistemic agent S believes that p is grounded in the fact that S believes that q), where p and q are the content of S's beliefs.

* Università Ca' Foscari Venezia, Italy
https://orcid.org/0000-0001-5858-7527
Department of Philosophy and Cultural Heritage, Università Ca' Foscari Venezia, Dorsoduro 3246, 30123 Venice, Italy
marco.sim@unive.it

Keywords: Aristotle's elenctic refutation; Epistemic dependence; Grounding; Law of Non-Contradiction; Moore, G.E.; Question-begging arguments.

1. Introduction

1.1. The Elenctic Strategy to Defend Both the Traditional Principle of Non-Contradiction and the Modern Law of Non-Contradiction

In this paper, I intend to show in what sense the elenctic strategy (elenchos, elenchus)1 to claim the truth of the Law of Non-Contradiction (hereinafter 'LNC') is not a fallacy of petitio principii. In doing so, I will also show why these considerations might reply to the dialetheist (i.e., the partial denier of LNC) without viciously begging the question (cf. §1.1 infra and §1.2 to understand the link between the alleged failure of elenchus and the consequential alleged success of dialetheism, as well as my reply through §§ 2.1-2.3).

Within the debate about the (necessary) truth of LNC, charging the elenctic strategy of being a vicious circularity is Costantini's (2018; 2020) ingenious objection, which, in turn, is based on the dialetheic account of negation developed by Graham Priest (1979), Priest (1998, 117-119). Indeed, according to Priest—who especially appeals to (Routley and Routley 1995)—'One may distinguish between three accounts of the relationship between negation, contradiction and content' (Priest 1998, 117).

The first account understands negation as cancellation: the operator 'not' deletes the content of the formula which is applied to.

---

[1] I use the phrases 'elenctic strategy', 'elenctic argument', 'elenchus', 'elenctic refutation', and the like as substantially equivalent. Further, unless otherwise stated, where I just mention the elenchus (and the like), I am referring to the elenctic refutation *of the denier of* LNC. Although the elenchus was used by Aristotle to defend his principle (cf. especially, *Metaphysics* IV.4, 1006a11-1006b34; and *infra* §1.1)—later called "Principle of Non-Contradiction" — for the sake of this paper I am going to assume that the key move of the elenctic strategy can also be invoked to defend the modern LNC ($\neg(\alpha \land \neg\alpha)$). I will return to this point below, within §1.1.

The second account, called ‘complementation account’, understands negation as in (modern) classical logic: the operator ‘not’ *always and only* excludes the content of what is denied (Costantini 2018). I will focus on this account later (see §1.2), since Costantini points out that such an account of negation, namely the classic account of negation, is the hidden assumption that turns the elenctic strategy into a vicious circularity. If we dropped out of the complementation account, the elenctic strategy in defence of LNC would not work (see §1.2). Moreover, I will read Costantini’s criticism against the elenchus especially through Bardon (2005), according to which there are *self-refuting* or *self-defeating* propositions, and among them the negation of LNC (*viz.* <LNC is false>),[2] but such a self-refutation takes place only if one holds certain *theoretical background assumptions* or *background presuppositions* (cf. §§.1.2-2.1). Combining (Bardon 2005) with Costantini’s objection, we will see that the putative self-refutation of a proposition like <LNC is false> would work only if we are prepared to assume the complementation account of negation.

The third account of negation belongs to paraconsistent logics, and it can be called ‘dialetheic’ account. Appealing especially to (Priest 1979), Costantini reads this account by leveraging the fact that negation *does not* always and only express exclusion: there are peculiar situations where the operator ‘not’ both excludes *and* accepts the formula which is applied to (see below §1.2).

Contra Costantini's objection against the elenchus, the aim of this article is to argue that there is at least one reading of the elenctic strategy in defence of LNC that *is not a vicious circularity*, *even assuming* the complementation account of negation. In doing so, I will leverage and adjust the distinction between two senses of epistemic dependence, offered by (Lemos 2004) and originally based on some thoughts about G.E. Moore’s ‘proof of an external world’ (1939; 1953)—see §2.2.

Indeed, my counter-objection aims to recover the elenchus in favour of LNC, *whilst maintaining a complementation account of negation.* In doing so, I will use a distinction between a necessary condition relation between propositions (p only if q) and a *grounding* relation between facts (the fact that an epistemic agent S believes that p is grounded in the fact that S

[2] Throughout the rest of this article, I use brackets <...> to indicate propositions.

believes that q)—where p and q are the content of S's beliefs, and, respectively, an argument's premise (for the sake of this paper: the elenctic argument for LNC) and its conclusion (see §§ 2.2-2.3). While the above-mentioned Lemos' distinction between two senses of epistemic dependence refers to the relations between propositions in both cases, I will propose to read the second sense of epistemic dependence in terms of grounding relations between facts, understood as a metaphysical and epistemic explanation (cf. §2.2). Such an adjustment can defuse Costantini's objection, leaving space to at least one reading of the elenchus in favour of LNC that is not viciously circular *and* assumes the complementation account of negation (cf. §2.3).

LNC states that no contradiction is true. By 'contradiction', I refer to either the *conjunction* between a proposition and its negation ($\alpha \wedge \neg\alpha$) or to the negation of the *identity* between a thing and itself: x: $x \neq x$. The latter turns out to be a sentence that denotes a (impossible) *contradictory* object, i.e., a non-self-identical thing (or entity), that is what Severino (1981, ch.4, §14) considers the *content* of a contradiction, namely nothing at all. Similarly, Oliver and Smiley (2013, 602) introduce a paradigmatic empty term, called '*zilch*', 'stipulating its impossibility of referring to something as a "logical necessity"'. Indeed, they formally define 'zilch' as x: $x \neq x$. Assuming that everything is self- identical, they conclude that 'zilch' is a term that necessarily fails to denote anything (cf. *ibidem*). We might say that 'zilch' picks up a contradictory object, i.e. a non-self-identical thing; but Oliver and Smiley—as well as Severino—do not accept contradictory objects in their ontology. Therefore, any term denoting a contradictory object is an empty term, i.e., a term that denotes nothing at all. According to Severino, the content of a contradiction is ultimately what results from a negation of the Law of Identity (insofar as, for unrestrictedly everything, x is an entity if and only if x is self-identical).[3]

[3] To sum up, when 'contradiction' refers to the conjunction between a proposition and its negation, we obtain the classic formulation of PNC: $\neg(\alpha \wedge \neg\alpha)$; when 'contradiction' refers to the negation of the identity between a thing and itself—or better: when we refer to the putative *content* of a contradiction -, *de facto*, we speak about (impossible, absolutely nonexisting) contradictory objects. In the latter case, the formulation of PNC might be something like: $\nexists x(x \neq x)$, where the domain of x is

Although the modern LNC differs from the Aristotelian 'principle'—later known as the 'Principle of Non-Contradiction' (hereinafter: 'PNC')—the classical treatment of a contradiction's negation can be traced back to Aristotle's works (Horn 2018), especially *Metaphysics* IV.3-6, e.g., *Metaph.* IV.4, 1005b19-22.[4] In their turn, both ontological and logical formulations of PNC '[are] traced in the writings of Parmenides, Gorgias, Plato' (Thom 1999, 153), with relevant differences and affinities pointed out by Thom (1999) but beyond the scope of this paper. In order to respond to Costantini's (2018; 2020) objection, according to which the elenctic method by Aristotle, and those who were inspired by it, especially Severino [1964] (1982; 2016), would be viciously question-begging (i.e., falling for a fallacy of *petitio principii*),[5] it is sufficient to consider the presence (explicit or implicit) of *negation* in both ontological and logical formulations of PNC, as well as in the modern LNC. As Thom (1999, 153) notes, 'The principle of non-contradiction received ontological formulations (in terms of "being" and "non-being") as well as logical formulations (in terms of affirmation and denial) in early Greek philosophy'. Now, as Costantini writes, 'What is essential to our ends is the presence of contradictory elements, and therefore of negation. [...] The whole game is played on the notion of negation' (2018, 850, translation mine). Indeed, the critical observation of Costantini on the elenctic defence of LNC focuses on the equivalence between *negation* and *exclusion* and on a certain way of understanding this (operation of) exclusion (cf. *infra*, §1.2). Sure, the Aristotelian conception of negation is different from modern (post-Fregean) ones, insofar as the former ranges primarily over *terms* and the latter over propositions and in any case never on sub-sentential units. Yet, both options are included under the so-called

unrestrictedly everything. A similar interpretation of PNC in terms of denying the existence of contradictory objects can be found in (Irwin, T. 1988).

4 Cf. Kirwan's translation, (Kirwan 1993[2], 7): 'For the same thing to hold good and not to hold good simultaneously of the same thing and in the same respect is impossible (given any further specifications which might be added against the dialectical difficulties).'

5 In the course of this paper, I use the phrases 'vicious circular argument', 'vicious circularity', '(fallacy of) *petitio principii*' (or simply: '*petitio principii*'), 'vicious question-begging (argument)', and the like as substantially equivalent.

'complementation' or 'classic' account of negation: cf. *supra* §1.1; *infra* §1.2; and (Priest 1998, 117 ff.). According to Costantini (2018), this account conceives negation only and always as *exclusion.* I would like to stress, together with Costantini, that what is at stake is not so much the object of the negation (either sub-sentential units, or propositions) but the negation *as such.*

It is also known that PNC has a special status for Aristotle, who claims it to be 'the firmest principle of all' (*Metaph.* IV.3, 1005b11-22). To him, PNC is not grounded in any hypothesis, being 'the principle of all the other axioms' (*Metaph.*, IV.3, 1005b32-34); it is the basis to build proofs and which in turn cannot be proved itself.[6]

We know, finally, that Aristotle proposes several strategies to defend PNC, *notwithstanding the impossibility of proving* it. According to Kirwan (1993[2]), in Aristotle's *Metaphysics* IV there are seven arguments in defense of PNC, but I will restrict my focus on the most known elenctic refutation (*elenktikos apodeixai*, also known as *elenchus* from Latin); therefore, the background of the following suggestions is *Metaphysics* (IV.4, 1006a11-1006b34).[7] Following, broadly speaking, the so-called Italian Neo-Scholasticism's general understanding of the elenchus,[8] I assume that the elenctic refutation consists in showing that, given a thesis, the negation of this thesis *implies* the thesis itself. In this regard, Pagani (1999, Part I, Ch. 2) points out that the relationship between the negation of PNC and PNC is not a relationship of *presupposition* but rather a relationship of *implication.* This important observation is also taken up by Costantini (2018; 2020), who applies it to the modern LNC as well, as we will see.[9] In the case of PNC,

[6] About the notion of *the firmest* principle and PNC as the firmest principle *of all*, cf. Wedin (2009, 133 ff.).

[7] Although the present article is not intended to be a commentary of Aristotle's works: cf. *infra* why and in which extent I appeal to Aristotle's elenctic refutation.

[8] Italian Neo-Scholasticism has been mainly developed around the Italian review *Rivista di Filosofia Neoscolastica* (founded in 1909, still existing: ISSN 0035-6247). Some scholars, either belonging to this tradition or coming from it, are mentioned across this article, like: Emanuele Severino; Sergio Galvan; Paolo Pagani. Bibliographical references are found across the text.

[9] Costantini (2018, 849 footnote, translation and emphasis mine) writes: 'What Pagani is saying here is that the denial of the Law of Non-Contradiction [LNC]—

the denier of PNC, that is, the one who intends to affirm the falsity of PNC, is forced (by the force of the *logos*, so to speak) to affirm the truth of PNC to the extent that she intends to say to herself and to others something that has at least one meaning (cf. Aristotle, *ibidem*). This strategy allegedly works both against she who claims that *all* contradictions are true (absolute negation of PNC) and against she who claims that *some* contradictions are true (partial negation of PNC). If the denier of PNC *actually* wants to declare precisely the negation of PNC *and not something else* (in either way), then—here is the elenctic refutation—she must in spite of herself affirm the truth of PNC. Otherwise, she would not deny PNC effectively: her negation would not be a negation, or would have no meaning, or she would be forced to remain silent, giving rise to no negation. The denial of PNC is therefore *self-refuting* (Bardon 2005 and cf. below §1.2), entailing a sort of self-negation (Severino [1964] 1982).

Now, I assume that the same Aristotelian elenctic strategy can be used to defend the modern LNC.[10] Indeed, following (Galvan 1995), (Pagani 1999), and (Costantini 2018), the key move of the elenctic strategy is the fact that the denial of LNC *necessarily implies* its truth; and—recall—I can switch from PNC to LNC because I have assumed—following (Costantini 2018)—that both of them *ultimately* share the view of negation as *exclusion.*

without the Law itself—would be not only self-contradictory, but even inconceivable. In this sense, the Law is a *condition of meaningfulness even for its own negation*'.

[10] See especially (Galvan 1995, 111): 'In the Aristotelean philosophical tradition, elenctic argumentation (*elenchus*) is conceived as a form of dialectical foundation of a thesis. It takes place in the context of discussion for and against a given thesis and consist in showing that, as the denier of this thesis argues against the opponent, he is unable to maintain his position unless he presupposes the thesis itself, which thus prevails and is consequently proven'. Here, Galvan uses the verb 'to presuppose', whilst Pagani (1999) and Costantini (2018)—and me, as well—insist on the fact that the key elenctic move is an implication. However, Galvan (1995, 112, emphasis added) himself states that a *stronger* application of the elenctic argument deals with implication: 'Elenctic argument is *more powerful when the implication between negation of the thesis and its assertion is necessary*; that is, when the opponent of the thesis in the end finds himself *necessarily obliged* to affirm it'. I will turn back to this key point in §2.1, Schema-ἐ, steps (2) and (3).

Therefore, *unless otherwise indicated*, from now on I will refer to the elenctic strategy as applied to the modern LNC.

The elenctic strategy was extended by Severino [1964] (1982; 2016), who, taking advantage of the defense inaugurated by Aristotle,[11] outlined two figures involved the elenchus. The first has as interlocutor, a hypothetical *absolute* denier of LNC. Meanwhile, the second is addressed to a supposed *partial* denier of LNC. Again, the absolute denier claims that LNC is always false, while the partial denier argues that there are situations in which LNC is false (or rather, as we will see in §1.2, situations in which LNC is both true and false—if she is a 'clever' denier). The two denials thus produced give rise to *trivialism* and *dialetheism*, respectively, two different philosophical positions according to which: '*Trivialism*: all contradictions are true (which implies that every proposition is true, since, for every proposition, we can consider its negation). *Dialetheism*: some contradictions (called 'dialetheias') are true' (Costantini 2018, 851, translation mine).[12]

The first figure of the elenchus shows that the absolute negation of LNC is *self-refuting* for the reasons already indicated above. To act as an absolute negation of LNC, the claim in question must mean *something*, precisely: the absolute negation of LNC and not something else (e.g., not the partial negation or the affirmation of LNC). Yet, in order for it to signify *something*, the absolute negation must confirm the truth of LNC (again, the negation of LNC *implies* LNC). The first figure of the elenchus, therefore, rules out trivialism.

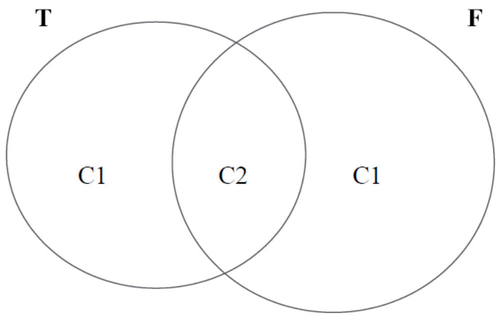
The second figure of the elenchus, which is more properly attributed to (Severino [1964] 1982), shows that even a partial denial of LNC is *self-refuting*. Presenting the prodromes of Severino's second elenctic figure, Priest (2020, 54-55, emphasis mine) writes:

---

[11] But perhaps a similar defense was already introduced by Socrates and then by Plato (Gottlieb 2023, in particular par.3 and par.9). It should be noted, however, that the elenchus differs from a *reductio ad absurdum*, the latter assuming the *impossibility* that contradictions are true, thereby already assuming LNC to be true (Perelda 2020, 13).

[12] To deepen the position of the dialetheist, in particular that of Graham Priest, in the relevant sense, I refer the reader to the bibliographical references quoted by Costantini (2018) and to the author himself. See especially the references given in (*ibidem*, 853, footnote).

Severino asks us to consider the following diagram:



> The lefthand circle contains those statements that are true; the right-hand circle contains those that are false (i.e. whose negations are true). The area of overlap is C2, *which contains things that are true and false*. The rest is C1. In the left part of this, things are true but not false; in the right, they are false but not true.

C2, therefore, is that *part* of language—so to speak—in which LNC is not true,[13] that is, where statements, propositions, or any *truth-bearer* is both true and false. This equals the *part* of reality where there are contradictory objects, i.e., non-self-identical things. Hence a *partial* denial of LNC, whereby the falsity of the law is attributed only to a *part* of language or reality (C2): *some* contradictions are true; *some* objects are contradictory. This is roughly the denial advanced by the dialetheist (although further clarifications are necessary; cf. §1.2). The elenctic method of (Severino [1964] 1982), in the case of C2, consists of pointing out that, for C2 to be the part of language or reality in which LNC does not apply, C2 must still respect LNC, that is, be *consistently itself and not* C1. The relation between C2 and C1, in short, also exemplifies a state of non-identity between two *different* positions (the partial denier of LNC *does not mean the same as the defender* of the absolute truth of LNC, *nor does she mean the same as the absolute denier* of LNC). However,

[13] We will see in §1.2, however, that, if the dialetheist were to describe the diagram, she would say that LNC in C2 is *true and is also false.* For further information, see Costantini (2018; 2020) and Priest (2020).

the non-identity between different positions is exactly an instance of LNC. So, even the partial denial of LNC is *self-refuting*: 'The first conclusion drawn from this is that the partial negation of LNC is self-contradictory' (Costantini 2018, 859, translation mine).

At this point, Severino grants a further chance to the partial denier of LNC, who is also—as we will see—doomed to failure (from Severino's point of view). Although C2 *as such*, that is, as a portion of language or reality, does not violate LNC (being a consistent part of language or reality), the *content* of C2 might be contradictory. In fact, the partial denier wants to affirm that there are (within C2) *truth-bearers* both true and false, or contradictory objects.[14] Among the examples of contradiction of the latter type, Severino ([1964] 1982) mentions the identification of two distinct items. For example, claiming the identity between the colour red and the colour green, attributable to the logical form <$x=y$>, i.e., <x is identical to y>, where indeed 'x' denotes the colour red and 'y' denotes the colour green.[15] Now, let us consider the identity between x and y. Severino distinguishes two interpretations to which the partial denier of LNC could allude.

In the first interpretation, 'x' and 'y' are two terms that both refer to *the same object*, for example, to a certain electromagnetic radiation of a certain wavelength, which—in the language used by the supposed denier—is referred to indifferently by 'red' and 'green'. In short, in this case the two

---

[14] With this strategy, Severino introduces a questionable theoretical assumption, namely, that we can quantify either on C2 in itself (the domain of dialetheias or contradictory objects) or on the content of C2 (the dialetheias or the contradictory objects). The first part of the disjunction (C2 in itself, i.e., C2 as a domain of quantification) is different from its members, therefore Severino formulates his elenctic strategy in the way we have just seen. However, Severino does not speak in terms of domains of quantification. Rather, he speaks (or would speak) in terms of parts of language or parts of reality. This exegetical and theoretical issue, however, can be overlooked as out of the scope of this paper.

[15] To be an effective identity between different terms, the term 'y' occurring in <$x=y$> is supposed to denote an object that is not identical to any object denoted by 'x'. So, <$x=y$> turns out to be a conjunction of <$x\neq x$> and <$y\neq y$>, where the contradiction is not due to the conjunction but to the negation of the Law of Identity in both conjuncts. Severino calls this (impossible) logical situation '*esser-diverso-da-sé*' (*being-different-from-itself*).

terms are *synonyms*, and it is evident that a denier of LNC is not producing an effective contradictory identity (if anything, she is using an anti-conventional use of the words 'red' and 'green'). In fact, given the reference to the same thing (the specific electromagnetic radiation of a certain wavelength), <x=y>, <red = green>, <the colour red *is identical to* the colour green> are all true propositions, with no problems in classical logic.

In the second interpretation, 'x' and 'y' refer to two *different* things: 'red' and 'green' refer to a single electromagnetic radiation of *two different* wavelengths (*at the same time* and *in the same respect*). Here, the use of the terms 'red' and 'green' is no longer bizarre; rather, the identification between the colour red and the colour green is what is bizarre, generating—precisely—a contradictory identity picking up a contradictory object. In this interpretation, the identity <x=y>, <*x is identical to y*> gives rise to an *authentic* contradictory identity. What has changed with respect to the first interpretation is that the two terms are *not synonyms*, that is, they do not refer to the same thing, but to two different things (namely, two electromagnetic radiations of different wavelengths), *despite them being identified.* And here Severino's trap is triggered: if x and y must be originally *different* (x must be itself and not y; y must be itself and not x; <red is identical to red>, <green is identical to green>, <x=y>, <y=y>) to finally denote a contradictory object, then the identity between x and *y is based* on their difference. Thus, the (partial) denial of LNC, exemplified by the proposition <x is identical to y>, is *self-refuting*, as it is based on the *difference* between x and y, which expresses exactly the deepest meaning of LNC according to Severino, that is, asserting the distinction between different items (and, conversely, the identity of what is self-identical). Explicitly or implicitly, the (partial) denier of LNC must affirm that x is not identical to y, <$x \neq y$>, when she really intends to refer to a *genuine* contradictory object, as opposed to appealing to a simple equivalence between synonyms that refer to the same self-identical object. If 'x' and 'y' are not synonyms, then the proposition <x=y> (<x is identical to y>) *is based* or *is grounded in* the proposition <$x \neq y$> (<x is different from y>, <x is not identical to y>). Talking about (relations of) *grounding*[16] is very useful for

[16] In §2.2 I will introduce and assume an account of grounding that might be fit for the sake of this paper: see also (Thompson 2019) and (Audi 2012). For

the purposes of this article and for comparison with what I will call the 'Moore-Lemos account' (cf. §2.2). In this regard, I will present an illuminating passage by Costantini (2020), which reconstructs the most important elenctic method of Severino, i.e., the last passage of the second figure of the elenchus in (Severino [1964] 1982, 49) in terms, exactly, of *grounding* (cf. §2.1).

1.2. The Objection (or Argument) by Costantini-Priest

In this section I reconstruct the argument by Costantini (2018; 2020) aimed at showing that the elenctic strategy in defence of LNC—analysed in the previous section—gives rise to a *petitio principii* (i.e., a vicious *question-begging* argument). Priest (1998; 2020) also raises a similar objection, or at least we can say that Costantini's objection is based on certain aspects of Priest's (1979; 1998).[17] Therefore, I will refer to these collectively as 'Costantini-Priest's argument' or 'Costantini-Priest's objection' or 'objection (or argument) by Costantini-Priest'. My counter-objection, proposed in §2.3, is mainly directed toward Costantini's formulation, but I believe that it *may* also be effective against Priest's (1998; 2020) under some respects, as both charge the Aristotelian elenchus of viciously begging the question. However, discussions of this hypothetical extension of my counter-objection are beyond the scope of this article.[18]

To reconstruct Costantini-Priest's argument against the elenchus, I use the concept of *self-stultifying proposition*, which we find in (Bardon 2005). According to Bardon (2005, 69 ff.), *self-refuting* or *self-defeating* propositions[19] are: (i) *self-referential* propositions, that is, they refer to themselves,

an overview of the notion of *grounding*, cf. (Bliss and Trogdon 2021) and (Raven 2015).

[17] On the link between Costantini's objection and the dialetheic account developed by Priest, cf. Costantini (2018, 849 footnote). See also *infra* §2.1.

[18] Costantini (2018; 2020) and Priest (2020, 49-59) mainly address the elenctic figures developed by Severino ([1964] 1982), based on the original Aristotelian strategy. There are, however, similar objections addressed *directly* to the Aristotelian defense: see especially (Priest 1998; 2020, 46-48).

[19] Bardon also deals with self-refuting *statements*. For the purposes of this article, I think I can overlook the distinction between propositions and statements, unless

to some aspect of the sentences that express them or to the performative acts (statements, utterances) of affirming them; *and* (ii) they can be expressed by *self-falsifying* statements, for example, the statement ‘I do not exist’ (*ibidem*, 70-71). Now, given the set of *self-refuting* propositions and statements, one might think that the denial of LNC (the proposition <LNC is false> or the statement that expresses it) falls within the typology of self-falsifying propositions or statements. But Bardon is keen to point out that self-refuting propositions should be divided into two subcategories: *self-falsifying* propositions, of the type just seen above, and *self-stultifying* propositions, which include – here’s the point – the denial of LNC. Bardon (2005, 73 ff., emphasis mine) distinguishes self-falsifying propositions from self-stultifying propositions as follows:

> Unlike a self-falsifying proposition, the [self-stultifying] proposition itself does not imply that its own affirmation should be impossible, and the affirmation of this proposition does not itself demonstrate that it is false. Rather, *what the proposition says or implies is inconsistent with one’s being epistemically entitled to affirm it.* [...] It is inconsistent to affirm a self-stultifying proposition because *that one is justified in making a claim is a pragmatic implication of making that claim.*

Among the examples of self-stultifying propositions indicated by Bardon (2005, 74), we find the denial of LNC (<LNC is false>).[20]

Why is it interesting and useful to start from here to reconstruct the objection by Costantini-Priest? It is because, in the definition of self-stultifying propositions (as well as in the definition of self-falsifying propositions), Bardon (2005, 73-74) clearly specifies that there must be *theoretical background assumptions* or *background presuppositions* for the ‘mechanism’ of

---

clearly necessary, as well as the distinctions between propositions, sentences, and utterances. In fact, my focus is more on propositions proper and on the belief in propositions than on their linguistic formulations, but, again, I do not think this is a topic of interest for the sake of the argument.

[20] Bardon uses ‘Principle of Non-Contradiction” (PNC), whilst I use ‘Law of Non-Contradiction’ (LNC) for the reasons I have already pointed out in §1.1.

self-*stultification* (as well as that of self-*falsification*) to take place.[21] Now, the defender of LNC (the one who appeals to the *self-stultification* of the proposition <LNC is false>) is accused of viciously begging the question by Costantini-Priest's objection *precisely because* – among the presuppositions or assumptions of her *theoretical background* – she holds '[...] *that account of negation* which is challenged by the friends of contradictions like Priest' (Costantini 2018, 849, abstract, emphasis mine).

At this point, to continue the exposition of Costantini-Priest's objection, it is necessary to identify what conception of negation occurs both in the assumptions of the defender of LNC and in the conclusion of the elenctic defense of LNC, that is, the conception that allegedly generates a vicious circularity. Costantini (2018, 854, translation and emphasis mine) identifies this conception in the classical meaning of negation as *exclusion*:

> [t]hose who deny LNC by claiming that there is at least one true contradiction are questioning the fact that *denial* is always able to *exclude* (the truth of) what is negated. *When you deny* LNC*, you are therefore denying the equivalence between negation and exclusion.*

What is challenged is that negation is *always and only* able to exclude what is negated. According to Costantini, this account of negation is, in fact, the one *theoretical background assumption* that the elenchus aims to ascertain as true. This classical account of negation is also known as the *complementation account*: cf. *infra* and (Priest 1998, 117 ff.).[22] Therefore, using that

---

21 It is interesting to note that Galvan (1995, 115, emphasis mine), in one of the most rigorous formalizations of the elenctic strategy, affirms: 'Elenctic argumentation *presupposes* the specification of a *common basis* of understanding between the denier of the thesis in question and its proponent'. This common basis is represented by a set of shared 'rules of logical deduction' (*ibidem*, 113 ff.) and 'a number of rules of *negation*' (*ibidem*, 114, emphasis added). *Mutatis mutandis*, in my reading of Costantini's (2018) treatment of the elenchus, I will understand the presence of the classic account of negation as the common ground shared by both the denier of LNC and the defender of LNC, where that common ground might be an exemplification of what Bardon (2005) calls 'theoretical background assumptions'.

22 Alongside the classic or complementation account there are at least two other accounts of negation: the so-called 'cancellation' account, according to which $\neg\alpha$

assumption to trigger the 'mechanism' of self-stultification of the proposition <LNC is false> (i.e., using that assumption *among the premises* of the elenctic argument) viciously *begs the question* (a point we will return to in §2.1).[23]

The complementation or classic account of negation can be expressed by the logical equivalence below:

(1) $T(\neg\alpha) \leftrightarrow \neg T(\alpha)$

---

cancels the content of α (Priest 1998, 117); and an 'intermediate' account from paraconsistent logics (*ibidem*), according to which 'the content of $\neg\alpha$ is a function of the content α, but neither of the previous kinds [namely, the complementation and the cancellation accounts]' (*ibidem*), as far as, for this account, a contradiction 'entails some things but not others' (*ibidem*). Besides, the complementation or classic account of negation is such that the content of a contradiction is *total* and 'entails everything' (*ibidem*), based on the *ex falso quodlibet* principle. Indeed, one of the main differences among the three accounts of negation—complementation or classical, cancellation, and paraconsistent accounts (especially the dialetheic one)—is linked to which content a contradiction ($\alpha \wedge \neg\alpha$) generates: respectively, everything, nothing, or something. However, as Priest (1998) notes, 'Though the cancellation and complementation accounts are quite distinct, some modern writers have run them together' (*ibidem*). I think that Emanuele Severino might be included among those writers, as far as he seems to use a classic account of negation, but, at the same time, he holds that the content of a contradiction is nothing at all. I leave this question open because it is beyond the scope of my paper. Furthermore, Severino's account of nothingness is more complex than what might seem (Severino 1981, ch. IV). However, about this specific topic, I just need to assume that a phrase like '*x: x≠x*' denotes nothing at all, like the empty term 'zilch' in (Oliver and Smiley 2013); cf. §1.1, regardless exegetical issues of Severino's works.

23 One could object (to Bardon and consequently to my way of introducing the argument by Costantini-Priest) that the denial of LNC is not a *self-stultifying* proposition but rather a *self-falsifying* one. Even if this were the case—and Bardon also contemplates this case, although he does not welcome it in (Bardon 2005, 90-91, footnote)—this would not compromise the key mark of the elenctic strategy that I intended to highlight in this section. Indeed, what interests me here is that, according to Bardon, to make both the *stultification* of a self-refuting proposition and the *falsification* of a self-refuting proposition work, theoretical *background assumptions* are needed.

where T is a truth predicate such as '…is true' or 'it is the case that…', and α is any truth-bearer (sentences, propositions, beliefs, etc.). Therefore, informally, (1) establishes that the negation of α is true if and only if α is not true. As Berto (2007, 6, emphasis mine) recalls, this idea 'expresses the semantics of *classical* negation, or the so-called *exclusion condition* of classical negation'.[24]

To complete the exposition of Costantini-Priest's objection, two other considerations are necessary. The first is that a true contradiction for the dialetheist, that is, a 'dialetheia', is not an *arbitrary* conjunction between contradictory propositions (the conjunction of a proposition and its negation). If so, we would be considering the position of the trivialist (cf. *supra*) and not that of the dialetheist. Costantini (2018, 862, translation and emphasis mine) explains this point very well:

> [accepting a contradiction, i.e. the conjunction of a proposition and its negation, as true] does not depend merely on the fact that [the dialetheist] wants to identify different items [or arbitrarily conjoin a proposition with its negation], as we can understand from an example of contradiction that Priest *does not accept*: I get on the bus and I don't get on the bus. […] Whenever it is asserted that '*x* is *y*' is a dialetheia *there must be a very specific reason that accounts for this assertion*. But this reason is not always present […].

[24] To be more accurate, (1) should be rendered by a sentence's (or another truth-bearer's) *name*, as Berto (2007) does: $T(\lceil\neg\alpha\rceil) \leftrightarrow \neg T(\lceil\alpha\rceil)$, where $\lceil\alpha\rceil$ is exactly the name of α. Furthermore, Berto correctly distinguishes $T(\lceil\neg\alpha\rceil) \leftrightarrow \neg T(\lceil\alpha\rceil)$ from $F(\lceil\alpha\rceil) \leftrightarrow T(\lceil\neg\alpha\rceil)$, namely, 'Sentence (or any truth-bearer) α is false if and only if its negation is true' (Berto 2007, 6). Although $T(\lceil\neg\alpha\rceil) \leftrightarrow \neg T(\lceil\alpha\rceil)$, namely, the equivalence between falsity and *untruth*, is more controversial than $F(\lceil\alpha\rceil) \leftrightarrow T(\lceil\neg\alpha\rceil)$, namely, the idea that 'false' means just '…has a true negation' (*ibidem*) (F being a falsity predicate), I will appeal to (1) when I refer to the classic or complementation account of negation throughout this paper, because, as Berto recalls, $T(\lceil\neg\alpha\rceil) \leftrightarrow \neg T(\lceil\alpha\rceil)$ expresses the exclusion condition of classical negation, that is exactly what Costantini (2018) points out as what makes the elenctic strategy for LNC a vicious question-begging argument.

There must be, therefore, a specific *reason* to affirm the truth of a contradiction, and that reason must be different from the mere willingness to contradict oneself or from the idea (naïve or not) that reality (or our representation of it) is contradictory. Indeed, as examples of dialetheias, Priest quotes logical or ontological scenarios in which, even if we try to deny the existence of contradictory objects or the conjunction of contradictory propositions (i.e., even if we try to apply LNC), we do not succeed (or rather, we succeed *and* not succeed; cf. *infra* and *ibidem*). We do not succeed because, in those specific logical or ontological situations, 'Negation fails to exclude the specific denied content' (Costantini, 862 footnote, translation mine). Priest's examples are well known in the scientific literature: the paradoxes of self-reference, transition states, paradoxes in set theory, borderline cases of vague predicates, etc.: see, e.g., (Priest and Berto and Weber 2022, par.3). Each of them defies LNC, that is, 'resists' the mere function of exclusion, thus showing that negation does not always and only express exclusion, that is, 'it does not work as expected by classical logic' (Costantini 2018, 855, translation mine).

The second consideration, useful for completing the exposition of Costantini-Priest's argument, consists of noting that 'the claim that there are true contradictions is not made from a consistent perspective. Rather, that very claim is a true contradiction' (Costantini, 855, translation mine). From the standpoint of the partial denier of LNC (the 'clever' or 'dialetheist' denier of LNC), even the partial negation of LNC is a dialetheia. So, the proposition <LNC is false> does not exhaust the content of the dialetheic negation of LNC, which *instead also* affirms the truth of LNC: <LNC is always true, but in some cases it is both true and false> (These are the aforementioned cases of logical or ontological scenarios in which the application of classical or complementation account of negation does not work because it fails to express *exclusion only.*)

2. How to Reply to the Dialetheist without (Viciously) Begging the Question

2.1. A Schema of Petitio Principii to Read Costantini-Priest's Objection

In this section, I propose an argumentative schema (I will call it 'Schema-ἐ'), of which the objection of Costantini-Priest to the elenctic strategy could be an example. Schema-ἐ will lay the ground for showing—in section §2.3—how one might reply to the partial denier of LNC without falling into a vicious circularity.[25]

The textbook definition of a question-begging argument is represented as follows:

$$\frac{\begin{matrix}\cdot\\\cdot\\ \mathrm{A}\\ \cdot\\ \cdot\end{matrix}}{\mathrm{A}}$$

That is, it is an argument that contains its conclusion among its premises. However, we might have a question-begging argument even though the conclusion—say *B*—was not identical to one of the premises—say *A*, where *A* entails *B*: see, e.g., (Iacona-Marconi 2005, 22 ff.).[26] At the same time, the textbook definition of *petitio principii* is controversial: I will come back to this topic at the end of §2.2.

Priest (1998; 2020) points out that Aristotle's elenctic defence of PNC or LNC (viciously) begs the question. For example, in (2020, 47, emphasis added), Priest writes the following:

---

[25] By this, I do not intend to exhaust all the possible schemas of *petitio principii* exemplified by Costantini-Priest's argument. However, I believe that it is more than sufficient to show (in §2.3) how to 'defuse' the charge of viciously begging the question.

[26] Coming from (Iacona-Marconi 2005, 20), the following argument, is an example of *petitio principii* where the conclusion is different from the premise:

$$\frac{\text{God created the Universe.}}{\text{God exists.}}$$

> Accepting that $\neg(A \wedge \neg A)$, or the stronger $\neg \Diamond (A \wedge \neg A)$, does *not* rule out accepting $(A \wedge \neg A)$. Of course, to do so is a contradiction. But *one cannot rule this out without supposing that one cannot accept a contradiction—which is exactly what is at issue* in disputes with the dialetheist.

Costantini (2018) argues that LNC is already one of the premises of the elenctic argument (therefore, making it a vicious question-begging argument) because the appeal to elenctic *refutation* is based on a *certain account of negation*, that is, the *classical negation* (or what has been called the 'complementation account', cf. *supra* §1.2). We can read Costantini's objection to the elenchus as a sort of focus on the reason why LNC is already assumed among the premises of the elenchus itself.[27]

As I anticipated in the previous section, let us indicate with '(1)' one of the premises of the elenctic refutation, specifically the above-mentioned classic account of negation, that is, the *theoretical background assumption* that negation *always* expresses *only* exclusion (cf. §1.2). Then, we can obtain the following schema for the elenctic strategy:

Schema-ἐ

(1) $T(\neg\alpha) \leftrightarrow \neg T(\alpha)$ [Assumption]

(2) $(\alpha \wedge \neg\alpha)$ [Assumption]

(3) $(\alpha \wedge \neg\alpha) \rightarrow \neg(\alpha \wedge \neg\alpha)$ [By self-refutation of (2)][28]

Therefore,

(4) $\neg(\alpha \wedge \neg\alpha)$ [2,3, Modus Ponens][29]

---

27 I think (although I am not sure) that Priest would agree with Costantini's criticism of the elenchus. Costantini's (2018; 2020) criticism is substantially based on (Priest 1979; 1998). Further, the charge of vicious question-begging assigned to the elenctic strategy already occurs in (Priest 1998) (although with several differences that are beyond the scope of this paper). Again, that is why I have chosen the term 'Costantini-Priest's objection' rather than simply 'Costantini's objection'.

28 With 'self-refutation' I refer to the idea by (Bardon 2005)—cf. §1.2—and the *implication* between contradiction and LNC (cf. §1.1).

29 To get the conclusion (4), one might alternatively appeal to the propositional *reductio* such that, if $p \vdash \neg p$, then $\vdash \neg p$. But this line of reasoning, which is essentially equivalent to the well-known *reductio ad absurdum*, already presupposes the truth of LNC, as noticed in §1.1, echoing (Perelda 2020, 13). Therefore, the use of

Assumption (1) represents the *complementation* account of negation, according to which $\neg\alpha$ has whatever content α does not have, i.e., α means *something different* from $\neg\alpha$ (Priest 1998, 117, and 2020, 52). The same assumption (1) can also be expressed as Costantini (2018, 857) claims: 'Negation is an operator behaving consistently', i.e., 'Negation always and only expresses (or means or implies) exclusion'.

Assumption (2) is what the denier of LNC intends to state. We need to assume (2) precisely because the elenctic strategy is supposed to be a defense against the *denier* of LNC.

The implication occurring in (3) is the core of the elenctic strategy. It is reasonable to infer (3), by self-refutation of (2), as far as the necessary condition of $(\alpha \wedge \neg\alpha)$ is the *difference* between what α and $\neg\alpha$ respectively mean. Indeed, if α meant the same as $\neg\alpha$, then their conjunction would not be a real contradiction. As we have already seen (cf. *supra* and §§1.1-1.2), the *complementation* account of negation reads negation always and only as *exclusion*, such that $\neg\alpha$ has whatever content α does not have. From the elenctic strategy standpoint, the negation of LNC (the antecedent of the implication occurring in (3)) *implies* LNC itself (the consequent), as I already pointed out (cf. §1.1). That means that who *in actu signato* claims any contradiction is *in actu exercito* denying the contradiction itself, therefore affirming the truth of LNC. In a nutshell, the denial of LNC is self- refuting. In §1.2, I accounted for this 'mechanism' of self-refutation following Bardon (2005, 73 ff.), who better clarifies this self-refutation in terms of self-stultifying propositions, whereby the *implication* between $(\alpha \wedge \neg\alpha)$ as antecedent and $\neg(\alpha \wedge \neg\alpha)$ as consequent could be *epistemically* understood.

---

that rule would not be fit to account for the elenctic strategy. Furthermore, I prefer to use *modus ponens* because I am fairly convinced that it is one of the most intuitive and universal rules of inference we can appeal to. Azzouni (2013, 3177) includes *modus ponens* (in its sentential version: [α and (α only if ß) only if ß]) in a set of logical steps and principles that 'any ordinary person will find intuitively unexceptionable'. Of course, someone could challenge them (and indeed it happened). Yet, if those principles are *introduced to an interlocutor in an appropriate manner,* then she/he should accept them (Azzouni 2013, 3178). Azzouni's standpoint looks even more interesting if compared to the elenctic defense of LNC, as far as he famously holds that natural language is logically inconsistent: see, e.g., (Azzouni 2013).

Conclusion (4) comes from *modus ponens*.

Let us focus again on the implication occurring in (3), the key step of the elenctic strategy. Appealing to Costantini's approach, one can object that the proponent of the elenchus affirms that $\neg(\alpha \land \neg\alpha)$ is the *necessary condition* of $(\alpha \land \neg\alpha)$ because she has *already assumed what she needs to prove,* i.e., *the complementation account of negation*, that is, that no proposition can be true and not-true (*untrue*) at the same time and in the same respect.[30] Indeed, the self-refutation of (2), resulting in the step (3), needs some *theoretical background assumptions* or *background presuppositions*, as Bardon (2005) notes about self-refuting propositions in general: cf. §1.2, where I proposed to read Costantini's objection by including the complementation account of negation—represented by (1) in my Schema-έ—among the theoretical background assumption of self-refutation.[31] What justifies the key step of the elenctic strategy is the idea that the necessary condition to hold a contradiction is the LNC itself. But the entire Schema-έ is viciously question-begging: assumption (1) and the conclusion (4) refer to the same idea. Generally, they say that *it is the case that* α is different from *it is the case that* $\neg\alpha$. In (1), this idea is expressed as a logical equivalence between *exclusion* and *negation* (cf. §1.2), whilst in (4) the same idea is expressed by denying the conjunction of α and $\neg\alpha$. Yet, both (1) and (4) somehow express what LNC essentially affirms, i.e., that $\neg\alpha$ *always and only excludes* α. If what premise (1) refers to is the same idea what conclusion (4) refers to, then the Schema-έ viciously begs the question. As Costantini (2018, 867-868, translation mine) says:

> [I]f it is already assumed [...] that negation always behaves only consistently [i.e., that *negation* always and only expresses or means *exclusion*], then the elenchus proves that there can be no true contradictions. Yet, if one wants to avoid such a *petitio principii* (for example by trying to prove exactly the assumption that negation always behaves only consistently), then the elenchus

30 I use the terms 'not-true' or 'untrue' due to the equivalence between falsity and untruth in the classic account of negation (cf. §1.2).

31 Therefore, step (3) also depends on (1), i.e., the exclusion condition of classical negation (cf. §1.2).

> cannot bring any additional contribution to the defense of LNC, which is not already present in LNC itself.

For a better understanding of Schema-ɛ́ and why it viciously begs the question, I would focus further on (3):

(3) $(\alpha \wedge \neg\alpha) \rightarrow \neg(\alpha \wedge \neg\alpha)$

Now, let us consider an instance of α, such that 'α' stands for $<x=y>$ and '$\neg\alpha$' stands for $<x\neq y>$, as far as, in light of the classic account of negation expressed by (1), 'it is true that *it is not the case that* $<x=y>$' is *logically equivalent to* 'it is not true that *it is the case that* x is identical to y.' As the reader will remember, $<x=y>$ or $<x$ is identical to $y>$ can be understood as an act of identifying two different items, ultimately referring to an (impossible) *contradictory* object x: $x\neq x$ (see §1.1).

Therefore, we obtain:

Schema-ɛ́ with 'α' standing for $<x=y>$

(1*) $T(<x\neq y>) \leftrightarrow \neg T(<x=y>)$ [Assumption]
(2*) $(<x=y> \wedge <x\neq y>)$ [Assumption][32]
(3*) $(<x=y> \wedge <x\neq y>) \rightarrow \neg(<x=y> \wedge <x\neq y>)$ [By self-refutation of (2*)]
Therefore
(4*) $\neg(<x=y> \wedge <x\neq y>)$ [(2*), (3*), *Modus Ponens*]

Let us consider the following notable excerpt by Costantini (2020, 102-103):

> The key point in Severino's argument is that the sentence '$x=y$' is an authentic negation of LNC only if x and y are not synonyms, i.e. only if '$x=y$' is *grounded* in '$x\neq y$'. In other words, to have a

[32] It might be interesting to note that two contradictions occur here. The first is due to the main conjunction. The second is 'internal' to the left conjunct because $<x=y>$ should be read as an identification of two different items, so that the left conjunct turns out to be $<x\neq x>$ and $<y\neq y>$, as I pointed out in §1.1, following (Severino [1964] 1982), and—to some extent— (Oliver and Smiley 2013). Of course, $<x\neq x>$ and $<y\neq y>$ are, in fact, two violations of the Law of Identity, namely, $\forall x(x=x)$, and their 'content' is given by (impossible) contradictory objects (non-self-identical things).

> contradiction, one must claim that x and y are distinct ($x \neq y$) and not distinct ($x=y$). The relation between the two contradictory sentences is one of grounding (‘$x \neq y$’ grounds ‘$x=y$’). This means that there is an asymmetry: ‘$x \neq y$’ may be true without ‘$x=y$’ being true, but not vice versa: in order to claim ‘$x=y$’ to be true (and to be an authentic negation of the LNC), the claim ‘$x \neq y$’ *must* be true too. The verb ‘must’ in the last sentence indicates that the truth of ‘$x \neq y$’ is a *necessary condition* for the truth of ‘$x=y$’. According to Severino, acknowledgement of the last point is enough to show that the denier of the LNC is wrong: her denial is grounded on what she is denying, and consequently the denial cannot be true.

Here, I want to anticipate and emphasize that—for the purposes of my paper—the relevance of this passage consists of its use of both the *grounding* relation and the *necessary condition* relation (*only if*). These are exactly the two points that my counter-objection will rely on (see §2.3). For the moment, though, let us just recall that, according to Severino, the act of identifying two different items *implies* the original difference of those two items (see §1.1). As Costantini (2020) correctly represents:

$<x=y> \rightarrow <x \neq y>$

To obtain this implication within the above application of Schema-ἐ, we just need to apply conjunction elimination to (2*) and then reiterate the self-refuting ‘mechanism’ for contradiction (already used in (3*)), assuming that $<x=y>$ is ultimately a contradiction such that x is not identical to x and y is not identical to y:

(5*) $<x=y>$ [(2*), conjunction elimination]

(6*) $<x=y> \rightarrow <x \neq y>$ [By self-refutation of (5*), assuming that $<x=y>$ is a contradiction *under specific conditions* (cf. §1.1)]

(7*) $<x \neq y>$ [(5*), (6*), *Modus Ponens*]

Following (Severino [1964] 1982) and (Costantini 2018; 2020), I read (5*) as a contradiction of the sort of $<x \neq x>$ (as well as $<y \neq y>$). Therefore, (6*) is exactly what Costantini points out as the core of Severino’s elenctic strategy. The antecedent occurring in (6*) is a way to deny LNC as far as ‘x’ and ‘y’ do not refer to the same object, yet they are *identified* (e.g., <the

color red is identical to the color green>). This identification can also be thought as a denial of the Law of Identity because, if y ‘picks up’ a different object (say, the color green) from what x denotes (say, the color red), then identifying x and y means affirming that x is not itself (e.g., <the color red is not identical to the color red>), since ‘y’ is supposed (by Severino) to denote an object that is not identical to any object denoted by ‘x’. In a nutshell, we can also think of the antecedent occurring in (6*) as <$x \neq x$> (and, *ceteris paribus*, <$y \neq y$>).

The consequent occurring in (6*) might be thought as an instance of LNC. Indeed, as we have seen in §1.1, according to Severino [1964] (1982), a way to express LNC consists in recognizing the difference of those items that are, *de facto*, thought of as different. In a nutshell, the essence of a contradiction is the *identity* between (or, better, the act of identifying) two different items that are originally thought of as *different*. That’s why Severino holds that the *difference* of any two *different* items is the *necessary condition* of any contradictory act of *identifying* them. This necessary condition relation can be exactly expressed by an *implication* between the identification of two different items (<$x=y$>) and their difference (<$x \neq y$>). As Costantini (2020, 103) points out, ‘$x=y$’ requires the truth of ‘$x \neq y$’, and ‘[t]he fact that ‘$x=y$’ requires the truth of ‘$x \neq y$’ implies that ‘$x=y$’ is simply false’ (*ibidem*). Yet, as Costantini (2020, 103, emphasis mine) notes,

> In classical logic, of two contradictory statements [*viz.* in our case ‘$x=y$’ and ‘$x \neq y$’] only one can be true. But *if negation is to be understood as classical*, then the argument is a *petitio principii*, because the dialetheist will argue that negation does not behave classically when dealing with true contradictions.

Therefore, following Costantini’s line of reasoning, we can conclude that Severino’s elenctic strategy against the existence of contradictory things (any object x such that $x \neq x$) is viciously question-begging, as far as the elenchus already assumes as true the complementation account of negation that shall be proved. Indeed, the ‘culprit’ of vicious circularity is that *theoretical background assumption*, i.e., (1) or—in this specific case—its exemplification (1*), conveying the classic account of negation, which (implicitly) is at work in (6*) in the form of a necessary condition relation for an authentic act of identification between two terms denoting two originally different items. As we have seen for the general Schema-ἔ, also in this specific instance of the schema *petitio principii* occurs: what (1*) refers to is essentially what (4*), namely, an exemplification of LNC, refers to. That is, broadly speaking, the idea that *it is the case that* $<x=y>$ is different from *it is the case that* $<x\neq y>$.

In the last section of the article, we will examine in detail how to prevent the elenctic strategy from raising to a *petitio principii* (namely, how to reply to a dialetheist without viciously begging the question). First, however, it is necessary to introduce (in §2.2) the account that I will apply to better understand the elenctic strategy (in §2.3), which consists of some thoughts by Moore (1953) and the relevant comments by Lemos (2004) about Moore's famous 'proof of an external world',[33] also charged with vicious circularity. I call this interpretive model the 'Moore-Lemos account'.

2.2. The Moore-Lemos Account and My Adjustments in Terms of Grounding

In response to those who charged Moore of vicious circularity for his 'proof of an external world', Moore ([1953] 1993, 77) writes:[34]

> Obviously, I cannot know *that* I know the pencil exits, unless I do know the pencil exists; and it might, therefore, be thought that the first proposition can only be mediately known—known *merely* because the second is known. But it is, I think, necessary to make a distinction. From the mere fact that I should not know

[33] As is known, Moore's argument for proving the existence of an external world goes as follows. I use the version that appears, e.g., in (Lemos 2004, 85):

Here is one hand.
Here is another hand.

Therefore, there are external objects.

Cf. (Moore 1939):

(1) Here are two hands.
(2) If hands exist, then there is an external world.
So there is an external world.

[34] In the quote, the external object is a pencil. In the best-known version, the external objects are Moore's own two hands.

> the first, *unless* I knew the second, it does not follow that I know the first *merely* because I know the second. And, in fact, I think I do know *both* of them immediately.

Setting the content of the argument about the external world aside, what I want to stress here is Moore's distinction between knowing a proposition *p* (which works as a premise) *only if* (*viz. unless*) you (already) know the proposition *q* (which works as a conclusion) and knowing that same proposition *p because* you (already) know *q*. According to Lemos (2004, 90, emphasis mine):

> Moore denies that the proposition 'S knows that *p only if* S knows that *q*' implies 'S knows that *p because* S knows that *q*'. From the fact that one knows that *p only if* one knows that *q* it does not follow that one knows that *p on the basis of* one's knowing that *q* or that *q* is one's *reason for* believing that *p*.

As we will see shortly, Lemos speaks both in terms of knowledge and in terms of belief. In fact, knowledge is traditionally treated as justified true *belief*. For the sake of this paper, then, I just need to consider belief. We therefore have the first tenet of what I call 'the Moore-Lemos account': one believes that *p* only if one believes that the proposition that *q* neither implies (*non sequitur*) (i) that the belief that *p* is based on the belief that *q*, nor (ii) that *q* is the reason why one believes that *p*. Regarding this tenet, one should keep in mind that the occurrence of '*only if*' exemplifies a necessary- condition relation, and that the occurrence of 'being *based on*' is equivalent to the use of 'because'. In a little bit, I will argue that the latter might exemplify a *grounding* relation, provided we introduce some appropriate adjustments (cf. *infra*). For the moment, though, let us focus on Lemos' account.

Lemos (2004, 90, emphasis mine) distinguishes two senses of epistemic dependence:

> Let us distinguish two senses in which one proposition can be 'epistemically dependent' on another. In the first sense, *p* is epistemically dependent$_1$ on *q* just in case one is justified in believing (or knows) *p only if* one is justified in believing (or knows) *q*. [...] But in a second sense, *p* is epistemically dependent$_2$ on *q* just in

> case one is justified in believing (or knows) p *on the basis of* one's being justified in believing (or knowing) q.

The fact that p is epistemically dependent$_1$ on q does not imply that p is epistemically dependent$_2$ on q. Let us clarify this difference with an example by Lemos (2004, 90) himself. Consider the argument, 'I think; therefore, someone thinks':

p = <I think>
q = <Someone thinks>
S is an epistemic agent who might believe p or q.

According to Lemos, the proposition <I think> is epistemically dependent$_1$ on the proposition <Someone thinks>. S believes <I think> *only if* S believes <Someone thinks>. Yet, the proposition <I think> is not epistemically dependent$_2$ on the proposition <Someone thinks>: S *does not* believe <I think> *because* she believes <Someone thinks>.

Another example might be extracted by the following argument: 'God created the Universe; therefore, God exists' (Iacona-Marconi 2005, 20). Applying Lemos' above-mentioned distinction, S believes <God created the Universe> *only if* she believes <God exists>: there is an epistemic dependence$_1$ relation between the premise and the conclusion of the argument. Yet—using Lemos' account—S *does not* believe that God created the Universe *because* she believes that God exists.[35]

To better understand this distinction, I think we need to introduce some adjustments to Lemos' (2004) account *in terms of grounding*. We will see that, in my reading, the epistemic dependence$_2$ does not hold between propositions (as does the epistemic dependence$_1$) but holds between (metaphysical and epistemic) facts. Hence, the epistemic dependence$_2$ becomes a kind of grounding relation. In doing so, I am going to change Lemos' conception of epistemic dependence$_2$ slightly but quite substantially. Let us see how. Indeed, neither Moore nor Lemos speak in terms of (metaphysical) grounding as the most recent literature does in the treatment of phrases, operators, or relations such as: 'because', 'in virtue of', 'on the basis of', and the like.

---

[35] Another example from Moore himself is exactly the perceptual knowledge that this is a pencil, which I have already recalled before (cf. *supra* §2.2.).

Before proceeding, I need to clarify which account of *grounding* might be suitable for the sake of this paper. As it is known, grounding is usually taken to be a 'a form of constitutive (as opposed to causal or probabilistic) *determination* or *explanation*' (Bliss and Trogdon 2021, introduction) between entities (e.g., facts). There are two broad understandings of grounding, according to one's attitude to either determination or explanation. Raven (2015, 326) calls them, respectively, 'separatism' and 'unionism', because the former *separates* grounding from metaphysical explanation, whilst the latter *unifies* them. Theorists of unionism, indeed, conceive grounding as a form of (metaphysical) explanation: <x grounds y> means <x explains y>. Theorists of separatism conceive grounding as a form of (metaphysical) determination: <x grounds y> means <x determines y>, namely, <x non-causally generates, produces, or brings about y>: see (Bliss and Trogdon 2021, §1.1); (Thompson 2019, 99-101). For the sake of this paper, I assume a unionist account of grounding, as far as Lemos' (2004) treatment of the original Moorean distinction (between 'because' and 'only if': see above §2.2) is explicitly *epistemic*, and the notion of (metaphysical) *explanation* seems to be exactly an epistemic affair as well (Thompson 2019, 101-103;). Moreover, I assume that grounding relations hold between facts, namely, obtaining states of affairs, rather than between truth-bearers (propositions, statements, or whatever).[36] As Raven (2015, 326) notes, 'Somehow, ground is *metaphysical* because it concerns the phenomena in the world itself, but also *explanatory* because it concerns how some phenomena hold in virtue of

36 For example, Audi's (2012) account of grounding establishes that the relation of grounding holds between facts, not between propositions (or other truth-bearers), where facts are what make propositions (or other truth-bearers) true. Audi's account belongs to so-called separatism because it understands grounding in terms of determination that backs explanation, whereas in this paper I have assumed a unionist approach. I think the reader might overlook this incongruity, since a fine-grained treatment of grounding is beyond the scope of this paper. Furthermore, we should remember that unionism and separatism might be intertwined if we conceive grounding as explanation (unionism) as *backed by* grounding as determination (separatism). However, as Bliss and Trogdon (2021, §1.1) notice, even if we agree that grounding is both *explanation* and *determination*, 'there still may be substantive reasons to go with one view rather than the other' (*ibidem*).

others.' Therefore, although I chose an epistemological approach to grounding to be closer to Lemos' (2004) reading of 'epistemic dependence$_2$', my choice could be compatible with a metaphysical approach, as far as the *epistemic* relation between facts is exactly a relation between *facts*, holding between *worldly* phenomena.

When speaking of grounding and explanation, this combination of metaphysics and epistemology is wisely treated by Thompson (2019). According to her, although metaphysical explanations concern worldly (objective) relations (in my assumption: relations between worldly *facts*), they should not be isolated by *our* (subjective) epistemic constraints: see especially (Thompson 2019, 101-103; 108). In particular, Thompson's (2019, 102, emphasis mine) approach to metaphysical explanation, namely, to what I assume grounding relations are,[37] introduces the above-mentioned epistemic constraints in forms of '*background beliefs* and *theoretical commitments* of the explanation seeker (and perhaps also of the explanation giver)'. The reader should note the relevant agreement between what I called *theoretical background assumptions* or *background presuppositions*, following (Bardon 2005)—see above §§1.2; 2.1—and what Thompson (2019) calls 'background beliefs and theoretical commitments.' For the sake of this paper, the most important theoretical commitment in question is the classic account of negation (see above §§1.2; 2.1). I will come back to this point later.

Finally, I assume that grounding relations are always (or almost always) transitive, irreflexive, and asymmetric.[38]

Provided with this account of grounding, or at least with these minimal *desiderata* for a hypothetical account of grounding, we can reinterpret Lemos' (2004) distinction between epistemic dependence$_1$ and epistemic dependence$_2$ as follows:

---

[37] Thompson (2019) does not make this assumption, developing her own account of metaphysical explanation regardless any particular view of grounding.

[38] There are other properties usually assigned to grounding relations (e.g. hyperintensionality, non-monotonicity, etc.) that are beyond the scope of this paper. Also, there are accounts of grounding relations that excludes such a relation to be irreflexive or asymmetric, for example. Again, these issues are beyond the scope of this paper.

(Epistemic dependence$_1$): the proposition *p* is epistemically dependent$_1$ on the proposition $q =_{\text{def}}$ *p* only if *q*, where *p*, *q* are the contents of S's beliefs.

(Epistemic dependence$_{2^*}$): the fact that S believes that *p* is epistemically dependent$_{2^*}$ on the fact that S believes that $q =_{\text{def}}$ the fact that S believes that *p* is grounded in the fact that S believes that *q*.

Propositions *p* and *q* are the contents of S's beliefs; the grounding relation occurring in the *definiens* of the epistemic dependence$_{2^*}$ should be read through the lens of the account of grounding assumed before. The reader should notice the relevant difference between the two epistemic relations: the epistemic dependence$_1$ is a relation that holds between *propositions* that are believed by an epistemic agent; instead, the epistemic dependence$_{2^*}$ is a relation that holds between *facts* (whilst Lemos' (2004) account conceives epistemic dependence$_2$ as a relation between propositions). In a nutshell, the epistemic dependence1 concerns a material implication between propositions, whilst the epistemic dependence$_{2^*}$ concerns a grounding relation (as metaphysical explanation) between facts (where the metaphysical explanation is at the same time epistemically constrained, since I partially assumed Thompson's (2019) account: see above).

In §2.3, I will apply these relations (epistemic dependence$_1$ and epistemic dependence$_{2^*}$) to our relevant case, namely, the elenctic strategy (as for Schema-έ).

The second tenet of the Moore-Lemos account is a *definition of petitio principii*: an argument is circular (in the vicious sense) if the belief in one of its premises is *based on* the belief in its conclusion. This definition[39] seems adequate to understand the basic idea of the vicious circularity argument exemplified by the Schema-έ of the elenchus as exactly a *petitio principii* (cf. §2.1). Following the Moorean distinction between 'only if' and 'because', or, better, the *non sequitur* already mentioned above, it is necessary to distin-

[39] Cf. Lemos (2004, 88-89, emphasis mine): 'Suppose we say that an argument begs the question if knowledge of a premise *is based* on knowledge of the conclusion'. Lemos here speaks in terms of knowledge but he immediately after speaks in terms of beliefs too (see *ibidem*, 90).

guish in turn between a) an argument whose logical form establishes a *necessary condition relation* between conclusion (q) and one of the premises (p), such that the necessary condition for believing that p is (already) believing that q; and b) an argument that has a logical form such that an *asymmetric relation* holds between the fact that an epistemic agent S believes the conclusion and the fact that S believes one of the premises: in other words, the belief that p is *based on* the belief that q. Lemos (2004) proposes that the argument of kind (b) is viciously question-begging, as opposed to that the argument of kind (a). Moore's 'proof of an external world' was discredited as being a *petitio principii* precisely because it was traced back to the argument of kind (b) by some of its critics (see *ibidem*). (In the next section, I will show how even the elenctic strategy—represented by Schema-ἐ—can avoid the charge of vicious circularity precisely because of this distinction between (a) and (b)). According to my adjustment of Lemos' (2004) epistemic dependence$_2$ in terms of a certain understanding of *grounding relations* (see above: epistemic dependence$_{2*}$), we might state that the argument of kind (b) is viciously circular as far as the fact that S believes the conclusion (q) *grounds* the fact that S believes one of the premises (p).

We assume (following Lemos) that epistemic dependence1 does not give rise to a *petitio principii*, whereas epistemic dependence2*, namely, my reading of Lemos' (2004) epistemic dependence2 in terms of grounding relations, does give rise to a *petitio principii.* [40]

In summary, the tenets that comprise my reading of Moore-Lemos account (hereinafter 'ML account' or just 'ML'), handy for the next section, are the following:

---

[40] Lemos (2004, 91) uses this assumption to defend Moore's 'proof of an external world'. For the sake of completeness, note that Lemos also hypothesises the objection that an argument could (viciously) beg the question even if one of the premises epistemically depended1 on the conclusion. Even then, he argues, Moore's 'proof of an external world' might *not* be a *petitio principii* (cf. *ibidem.*). However, here I do assume that an argument viciously begs the question when a grounding (asymmetric) relation holds between the *fact* that an epistemic agent believes one of the premises and the *fact* that the very same epistemic agent believes the conclusion, whereas there is *no petitio principii* when the relationship between the conclusion and premise is a necessary condition relation between propositions (p only if q).

(ML1) If the premise of an argument is epistemically dependent on the conclusion of that argument, then the epistemic dependence can be either a *necessary condition relation* between *propositions* (understood as contents of beliefs), or a *grounding relation* between *facts*.

(ML2) From the fact that one believes the proposition that *p* only if (*necessary condition relation*) one believes the proposition that *q*, it *does not* follow that the fact that an epistemic agent S believes that *p is grounded* in the fact that S believes that *q*.

(ML3) An argument in which one of its premises *p* is epistemically dependent on its conclusion *q* is viciously circular (i.e., a *petitio principii*) when the epistemic dependence exemplifies a *grounding* (asymmetric) relation between facts (what I have called 'epistemic dependence$_{2*}$') but not when the epistemic dependence exemplifies a *necessary condition relation* between propositions (what Lemos calls 'epistemic dependence$_1$).

Before moving forward, it is worth considering the relationship between valid arguments and instances of *petitio principii.* As for example Iacona and Marconi (2005) point out, the philosophical literature does not undisputably place the border between *valid* arguments and *invalid* arguments in the case of question-begging arguments. Indeed, 'Although it is uncontroversial that there is something wrong with begging the question, it is not clear from those definitions *what* is wrong' (Iacona and Marconi, 2005, 19). Since the ML account deals with *petitio principii* in terms of *epistemic* dependence, I assume that question-begging arguments should be assessed *epistemically*, as Lemos (2004) does, and in accordance with my above reading in terms of grounding whereby the relation of grounding is both a sort of metaphysical and *epistemological* explanation. Iacona and Marconi (2005) clearly summarize this kind of approach into *petitio principii*, originally based on (Sanford 1972), as follows (although they propose a different approach in the *pars construens* of their article):

> According to a rather popular line of thought […] begging the question is to be defined in terms of some epistemic relation between one or more premises and the conclusion. One way of putting things consists in saying that the relation involves the actual beliefs of the person to whom the

> argument is addressed. In this vein, a question-begging argument may be defined as an argument addressed to someone who believes one or more of the premises *only because* he already believes the conclusion, or to someone that would believe one or more of the premises *only if* he already believed the conclusion (Iacona and Marconi 2005, 25, emphasis mine).

About this definition, it is worth underlining that both a sort of epistemic grounding relation ('[…] only because […]') and a necessary condition relation ('[…] only if […]') are mentioned: the reader can easily note that these may be those kinds of epistemic dependence relations that we have found in the ML account, and especially in my reading of Lemos' account (in my reading: epistemic dependence$_{2*}$, rather than Lemos' own epistemic dependence$_2$). This parallels the claim that an argument begs the question when the epistemic dependence exemplifies a *grounding* (asymmetric) relation but not a *necessary condition* relation (ML3). For the sake of my argument, this is a relevant difference between Sanford's definition of (putative vicious) question-begging arguments (where *grounding* or *necessary-condition* relations between a premise and a conclusion might generate a *petitio principii*) on the one hand, and both the original ML' definition and my reading of it (where only *grounding* might generate a *petitio principii*) on the other hand.

2.3. A Reply to the Partial Denier of the Law of Non-Contradiction

In this section, I will apply the ML account to reinterpret Schema-έ (occurring in §2.1) which expresses Costantini-Priest's objection, that is, the thesis that the elenctic strategy is a *petitio principii.* Using the ML account, we will see in what sense the elenctic refutation of LNC's denier does not give rise to a *petitio principii.* This means providing a non-question-begging reply to the denier of LNC.

Let us recall Schema-έ:

(1) $T(\neg\alpha) \leftrightarrow \neg T(\alpha)$ [Assumption]
(2) $(\alpha \wedge \neg\alpha)$ [Assumption]
(3) $(\alpha \wedge \neg\alpha) \rightarrow \neg(\alpha \wedge \neg\alpha)$ [By self-refutation of (2)]

Therefore,

(4) $\neg(\alpha \land \neg\alpha)$ [(2), (3), *Modus Ponens*]

We can apply the ML account to read Schema-ἐ, focusing on the epistemic relation that holds between the premise (1), i.e., the classic account of negation, and the conclusion (4), i.e., LNC, according to which there are no true contradictions. (For easier reading, consider that, here, (1) represents the premise *p*, and (4) represents the conclusion *q* of the general explanation of the ML account). Indeed—as I pointed out in §2.1—the (putative) *petitio principii* occurs because the elenctic defender of LNC *already assumes* the conclusion (4), i.e., LNC itself, in order to believe (or understand) the premise (1), i.e., the classic account of negation. Now, recalling ML1, ML2, and ML3 together with the propositions of Schema-ἐ, my argument to defuse *petitio principii* accusation runs as follows:

(A1) Premise (1) is epistemically dependent on the conclusion (4)

(A2) Given an epistemic agent S, the epistemic dependence relation occurring in (A1) can be read either as a necessary condition relation between propositions that are believed by S ((1) *only if* (4)), or as a grounding relation (the fact that S believes (1) *is grounded in* the fact that S believes (4))

(A3) Premise (1) is true *only if* the conclusion (4) is true.

(A4) From the fact that S believes premise (1) *only if* S (already) or believes the conclusion (4), *it does not follow* that S believes the premise (1) *because* S believes the conclusion (4).

Therefore,

(A5) Schema-ἐ does not viciously beg the question.

Let us assess this line of reasoning. According to the ML account, the epistemic dependence$_1$ does not give rise to any *petitio principii*. (A1) is my starting point as far as I need to reply to Costantini- Priest's objection that the elenctic strategy viciously begs the question. In fact, I concede that there is an epistemic dependence between (1) and (4). (A2) is obtained by applying (ML1) to (1) and (4). (A3) represents how I mean to read the epistemic dependence between (1) and (4) in Schema-ἐ: negation always and only expresses exclusion, or negation is an operator that behaves consistently, *only if* there are no true contradictions. In a nutshell, believing premise (1) *epistemically depends$_1$* on the conclusion (4).[41] (A4) is obtained by applying (ML2) to (1) and (4). Let us check how. Let us consider the following:

i. The propositional content of (1) is <The complementation account of negation is true>.
ii. The propositional content of (4) is <LNC is true>.
iii. S is an epistemic agent who might believe propositions p (premise of Schema-έ) or q (conclusion of Schema-έ)

By (ML2), from the fact that S believes (1) *only if* she believes (4) *it does not follow* that S believes (1) *because* she believes (4), where the operator 'only if' can be read as an epistemic dependence$_1$, whilst the operator 'because' can be read as an epistemic dependence$_{2*}$. In other words, the fact that the truth of LNC is the necessary condition of the truth of the classic account of negation *does not entail* that LNC (metaphysically and epistemically) grounds the classic account of negation. The *rationale* of conclusion (A5) is (ML3).

Schema-έ would indeed give rise to a vicious circularity if we replaced assumption (A3) with the following (A3*):

(A3*) Premise (1) is true *because* the conclusion (4) is true, namely, the fact that S believes premise (1) is grounded in the fact that S believes the conclusion (4).

In this case, we would obtain an epistemic dependence$_{2*}$ between (1) and (4). Consequently, by (ML3):

(A5*) Schema-έ viciously begs the question.

So, whilst the latter reading of Schema-έ gives rise to a *petitio principii*, the former reading—{A1; A2; A3; A4; A5}—does not viciously beg the question.

41 About the difference between epistemic dependence$_1$, epistemic dependence$_2$, and epistemic dependence$_{2*}$, see §2.2.

Similar considerations can also be made when 'α' stands for $<x=y>$, namely, when our focus is on the act of identifying different items (x, y) or—in a nutshell—when we refer to contradictory objects (non-self-identical things) such as $x \neq x$ (cf. §1.2). In this case our focus is on what Severino (2005, *passim*) calls 'the content of a contradiction', namely, the identity of different items that, *de facto*, turn out to be a contradictory object. As Costantini (2018; 2020) effectively highlights, the core of Severino's elenctic strategy is represented by the implication below (cf. §1.2 and §2.1):

(6*) $<x=y> \rightarrow <x \neq y>$

If we compare Costantini's reconstruction of Severino's elenctic strategy (Costantini 2020, 102- 103) with the ML account, we immediately notice that both the *necessary condition* relation (*only if*) and the *grounding* relation appear in it. It seems to me, however, that these two kinds of relation are not properly separated in his argument, as Costantini uses the '*id est*' (*ibidem*, 102) just to explain that the necessary condition relation resolves into a *grounding* relation between the two sentences—'$x=y$' is *grounded* in '$x \neq y$'—or between the two related propositions, or, again, according to my reading of the ML account, between the fact that an epistemic agent S believes one proposition and the fact that S (already) believes the other. Meanwhile, the ML account invites us to distinguish the two relations within a given argument (see ML1 and ML2). If we apply the ML account, especially the distinction between epistemic dependence$_1$ and epistemic dependence$_{2*}$, in reading the reconstruction by Costantini (2020) of Severino's elenctic strategy, then we have the following:

(C1) Where Costantini (2020, 102, emphasis mine) writes 'the sentence '$x=y$' is an authentic negation of LNC *only if* x and y are not synonyms', we can understand this to mean that the proposition $<x=y>$ is *epistemically dependent$_1$* on the proposition $<x \neq y>$ (notwithstanding the fact that the identity between x and y must be understood as an effective or authentic contradiction). Therefore, $<x=y> \rightarrow <x \neq y>$, where $<x \neq y>$ is a *necessary condition* of $<x=y>$, as Costantini also observes.

(C2) Where Costantini writes that '$x=y$' is *grounded* in '$x \neq y$' (*ibidem*), we can understand this to mean that the fact that S believes the proposition $<x=y>$ is *epistemically dependent*$_{2^*}$ on the fact that S believes the proposition $<x\neq y>$.

(C3) Where Costantini (2020, 102-103) writes (or at least suggests) that $<x=y>$ *is grounded in* $<x\neq y>$, we can understand this to mean that S believes that $<x=y>$ *because* or *on the basis of* her belief of the proposition $<x\neq y>$. In other words, the partial denier of LNC must *already be aware* of the difference between x and y, namely, the terms she wants to identify in an effective (real, authentic, and true) contradiction.

(C4) Where Costantini (*ibidem*) writes (or at least suggests) that $<x\neq y>$ *grounds* $<x=y>$, we can understand this to mean that the fact that S believes $<x\neq y>$ is *the metaphysical and epistemic explanation* of the fact that S believes that $<x=y>$.

If we read the epistemic relation between the proposition $<x=y>$ and the proposition $<x\neq y>$ in terms of epistemic dependence$_1$ (as it occurs in (C1)), then the version of Schema-έ with 'α' standing for $<x=y>$ does not viciously beg the question (by the ML account). Instead, if we read the same epistemic relation in terms of epistemic dependence$_{2^*}$—as it occurs in (C2)—then that version does give rise to a *petitio principii* (by the ML account). Again, there is at least one reading of Schema-έ that does not involve any vicious question-begging also when 'α' stands for $<x=y>$, i.e., when our focus is on a proposition that—so to say—describes an (impossible) fact: the fact that there is a putative contradictory object x such that $x\neq x$ (cf. §1.1).

Note that my interpretation of the 'heart' of the elenctic strategy *does not challenge* the *core* of Costantini-Priest's objection, according to which the elenchus presupposes the conception of 'classical negation' (see assumptions (1) and (1*)), which the dialetheist does not assume and, indeed, questions. My counter-objection to Costantini-Priest's objection, in effect, only concerns the charge of *petitio principii*. That is, even accepting that the classic account of negation as (always and only) *exclusion* is a *theoretical background assumption* of the elenctic strategy, it does not mean that the elenctic strategy viciously begs the question. This is because—given ML—the game on the charge of *petitio principii* is played on the difference between epistemic dependence$_1$ and epistemic dependence$_{2^*}$, namely, two different epistemic relations we can use to understand the elenchus in one way or another (either in a way that *does not* viciously beg the question or in a way that *does*).

2.4. Concluding Remarks

Let us return to Schema-ἔ and my reading of it ({A1; A2; A3; A4; A5}) according to which the elenctic strategy does not viciously beg the question. The focus of that reading is on the epistemic relationship between premise (1) and the conclusion (4). Since (1) expresses the complementation account of negation (cf. §1.2), and (4) is the propositional formulation of LNC, then my *epistemic* interpretation ({A1; A2; A3; A4; A5}) of the elenctic strategy affirms that *already* believing in LNC is the *necessary condition* of believing in the complementation account of negation. That is, S believes (1) *only if* S already believes (4). As we have seen (§§2.2-2.3), according to my reading (based on the ML account), the epistemic dependence of premise (1) on the conclusion (4) does not give rise to a *petitio principii.* If the epistemic dependence of (1) on (4) were understood in terms of *grounding*, then the epistemic dependence$_{2^*}$ would give rise to a *petitio principii.* Now, what about the grounding relation? We already know that—according to my reading ({A1; A2; A3; A4; A5})—the fact that S believes LNC (4) *does not ground* the fact that S believes the classic account of negation (1), by selecting (A3) rather than (A3*). However, could we still somehow or somewhere admit a grounding relation? I would answer this question in two ways.

(i) We might claim that our belief in the classic account of negation, i.e. (1), *grounds* our belief in LNC, i.e., (4). That means that the fact we believe the conclusion of the elenctic strategy (4) *is grounded* (at least) in the fact that we believe in one of the premises of the elenctic strategy (1). Since the ML account does not prevent us from accepting a grounding relation where the fact that S believes the *conclusion* of an argument *is based* on the fact the S believes in one of the premises of that argument, it might be the case that S holds a grounding relation between the same two epistemic facts but running the opposite way, rather than the way of epistemic dependence$_{2*}$, without viciously begging the question. In a nutshell, this option rejects the grounding relation expressed by the epistemic dependence$_{2*}$, namely the idea that believing the premise of an argument is based on (already) believing the conclusion of that argument, but such an option keeps the grounding relation by *reversing* it, namely, accepting that believing the conclusion of an argument *is based* on believing the premises of that argument. In fact, it seems intuitive and plausible to maintain that—broadly speaking—the conclusion of an argument immune from charges of vicious circularity is based on its own premises (whilst one of the premises of a vicious *question-begging* argument would be based on its own conclusion).

(ii) Alternatively, we might 'remain silent' on any kind of grounding relation between the premises and the conclusion of the elenctic strategy, appealing just to necessary condition relation between propositions (epistemic dependence$_1$, 'only if').

Although I think both options are available to my argument, I would prefer the first option (i). Indeed, the first option can account for Costantini's (2018) idea that LNC *presupposes* the classic account of negation. However, Costantini reads that presupposition without distinguishing the kind and the direction of the epistemic dependence between the classic account of negation and LNC, thereby charging the elenchus with a *petitio principii*. My proposal, combining my argument ({A1; A2; A3; A4; A5}) with option (i), reads this claim exactly disambiguating the kind and direction of that epistemic dependence.

I would conclude with another remark on *grounding*, this time between LNC and the *overall* elenctic strategy. Costantini (2018, 17) knows well that (for Severino, and earlier for Aristotle) LNC is not grounded in the elenchus itself. Rather, the elenctic refutation limits itself to *ascertaining* the truth of LNC.[42] However, my objection to Costantini (exploiting the

[42] As Perelda (2020, 14) writes, 'This argument [*viz.* elenchus], mind you, does not *ground* the principle [*viz. LNC*]: it is not a *reason* for the *truth* of the principle which has none (if *reason* means something that grounds the truth of the principle)'.

ML account) is the lack of a sufficient distinction between grounding relations and the necessary condition relations in his treatment, as we have seen. My objection might also extend to the way in which Severino accounts for the relation between contradictory items (both truth-bearers and objects) and non-contradictory items, insofar as Severino implicitly—or explicitly—suggests that it is a grounding relation. To Severino, though, this relation does not hold between LNC and the elenchus but between the negation of LNC (i.e., asserting the identity of different items) and LNC itself (i.e., asserting the distinction between different items and, *ceteris paribus*, the identity of what is self-identical). The present article, however, had no exegetical purposes regarding Severino's works. So, the fact that the *grounding* relation is an excellent way of paraphrasing what Severino claimed (and that therefore (Costantini 2020) too provides an excellent commentary on Severino's theses) does not mean that this is also an adequate way to do justice to the elenchus.

Acknowledgements

I would like to thank Filippo Costantini, Graham Priest, and an anonymous referee for this journal for their helpful feedback. I also thank the attendees of the conference "Aristotle across Boundaries" (Verbania, 8 June 2023), organized by Silvia Fazzo (Università del Piemonte Orientale), during which an early draft of this paper was presented.

References

Audi, Paul. 2012. "Grounding: Toward a Theory of the 'In-Virtue-Of' Relation." *The Journal of Philosophy*, 109(12): 685–711. https://doi.org/10.5840/jphil20121091232

Azzouni, Jody. 2013. "Inconsistency in Natural Languages." *Synthese*, 190(15): 3175–184. https://doi.org/10.1007/s11229-012-0136-5

Bardon, Adrian. 2005. "Performative Transcendental Arguments." *Philosophia*, 33(1): 69-95. http://dx.doi.org/10.1007/BF02652648

Berto, Francesco. 2007. *How to Sell a Contradiction. The Logic and Metaphysics of Inconsistency.* London: College Publications.

Bliss, Ricki, and Trogdon, Kelly. 2021. "Metaphysical Grounding." *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition), edited by Edward N. Zalta. https://plato.stanford.edu/archives/win2021/entries/grounding/

Costantini, Filippo. 2018. "Elenchos come *petitio principii.*" *Rivista di Filosofia Neo-Scolastica*, 110(4): 849–70.

Costantini, Filippo. 2020. "Challenging the Law of Non-Contradiction." *Eternity & Contradiction. Journal of Fundamental Ontology*, 2(2): 90-107.

Galvan, Sergio. 1995, "A Formalization of Elenctic Argumentation," *Erkenntnis*, 43(1): 111–26. http://dx.doi.org/10.1007/BF01131842

Gottlieb, Paula. 2023. "Aristotle on Non-contradiction." *The Stanford Encyclopedia of Philosophy* (Winter 2023 Edition), edited by Edward N. Zalta and Uri Nodelman. https://plato.stanford.edu/archives/win2023/entries/aristotle-non-contradiction/

Horn, Laurence, R. 2018. "Contradiction." *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition), edited by Edward N. Zalta. https://plato.stanford.edu/archives/win2018/entries/contradiction/

Iacona, Andrea and Marconi, Diego. 2005. "Petitio principii: What's wrong?" *Facta Philosophica*, 7(1): 19–34. http://dx.doi.org/10.5840/10.3726/93519_19

Irwin, Terence. 1988. *Aristotle's First Principles*, section 98. Oxford: Oxford University Press.

Kirwan, Christopher, ed. 1993[2]. *Aristotle; Metaphysics, Books* Γ*,* Δ*, E.* Oxford: Oxford University Press.

Lemos, Noah. 2004. *Common Sense. A Contemporary Defense*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511498800

Moore, George Edward. 1939. "Proof of an External World." *Proceedings of the British Academy*, 25(5): 273–300. Reprinted in Baldwin, Thomas. 1993 ed. *Selected Writings*, 147-170. London and New York: Routledge. https://doi.org/10.4324/9780203769812

Moore, George Edward. 1953. "Hume's Theory Examined." In *Some Main Problems of Philosophy*, 108-126. London: George Allen and Unwin. Reprinted in Baldwin, Thomas. 1993 ed. *Selected Writings*, 59–78. London and New York: Routledge. https://doi.org/10.4324/9780203769812

Oliver, Alex and Smiley, Timothy. 2013. "Zilch." *Analysis*, 73(4): 601-613. http://dx.doi.org/10.1093/analys/ant074

Pagani, Paolo. 1999. *Contraddizione performativa e ontologia.* Milano: Franco Angeli.

Pagani, Paolo. 2020. "La forma dell'*Élenchos.*" *Rivista di Filosofia Neo-Scolastica*, 112(4): 969–94.

Perelda, Federico. 2020. "Who is Afraid of Contradictions? A General Introduction to the Debate Between Severino and Priest." *Eternity & Contradiction. Journal of Fundamental Ontology*, 2(2): 6–30.

Priest, Graham. 1979. "The Logic of Paradox." *Journal of Philosophical Logic*, 8(1): 219-241.

Priest, Graham. 1998. "To Be *And* Not To Be—That is the Answer: On Aristotle on the Law of Non- Contradiction." *History of Philosophy & Logical Analysis*, 1(1): 91–130

Priest, Graham. (2020). "Emanuele Severino and the Principle of Non-Contradiction." *Eternity & Contradiction. Journal of Fundamental Ontology*, 2(2): 42-66.

Priest, Graham, and Berto, Francesco, and Weber, Zach. (2022), "Dialetheism." *The Stanford Encyclopedia of Philosophy* (Fall 2022 Edition), edited by Edward N. Zalta and Uri Nodelman. https://plato.stanford.edu/archives/fall2022/entries/dialetheism

Raven, Michael, J. 2015. "Ground." *Philosophy Compass*, 10(5): 322–33. https://doi.org/10.1111/phc3.12220

Routley, Richard and Routley, Valerie. 1995. "Negation and Contradiction." *Rivista Colombiana de Matematicas*, 19(1-2): 201–31.

Sanford, David, H. 1972. "Begging the question." *Analysis*, 32(6): 1971-99.

Severino, Emanuele. [1964] 1982. "Ritornare a Parmenide." *Rivista di Filosofia Neo-Scolastica*, 56(2): 137–75. Reprint in *Essenza del nichilismo*, 19-61. Milano: Adelphi.

Severino, Emanuele. 1981. *La struttura originaria*. Milano: Adelphi

Severino, Emanuele. 2005. *Fondamento della contraddizione*. Milano: Adelphi.

Severino, Emanuele. 2016. "Returning to Parmenides." In *Essence of Nihilism*, 35–83. Translated by Giacomo Donis. London and New York: Verso.

Thom, Paul. 1999. "The Principle of Non-Contradiction in Early Greek Philosophy." *APEIRON. A Journal for ancient philosophy and science*, 32(3): 153–70. http://dx.doi.org/10.1515/APEIRON.1999.32.3.153

Thompson, Naomi. (2019). "Questions and Answers: Metaphysical Explanation and the Structure of Reality." *American Philosophical Association*, 5(1): 98–116. http://dx.doi.org/10.1017/apa.2018.51

Wedin, Michael V. 2009. "The Science and Axioms of Being." In *A Companion to Aristotle*, edited by G. Anagnostopoulos, 123-143. Oxford: Wiley-Blackwell. https://doi.org/10.1002/9781444305661.ch8

Organon F 31 (2) 2024: 183–196
https://doi.org/10.31577/orgf.2024.31204
ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Cognition As a Natural Kind

Selen Fettahoğlu*

Received: 30 October 2023 / Revised: 22 May 2024 / Accepted: 25 May 2024

Abstract: In corvids and apes, cognition evolved convergently instead of being inherited by a shared ancestor. In biology, natural kinds are classified according to common ancestry. So, if we were to apply the same strategy to psychology, cognition among corvids and apes would not be the same natural kind. However, Cameron Buckner claims that cognition is a natural kind. I suggest that by using Ladyman and Ross's strategy of taking natural kinds as real patterns, we can support that cognition is a natural kind. Cognition seems to have the properties of predictability and compressibility, which are necessary conditions for real patternhood. Thus, convergent evolution examples of cognition, such as that found in corvids and apes, can be the same natural kind.

Keywords: Natural kinds; cognition; convergent evolution; real patterns.

1. Introduction

Assuming that there are natural kinds allows us to track regularities in science. To speak about natural laws in science, which is a common practice

* Ibn Haldun University; Middle East Technical University
https://orcid.org/0000-0002-0749-052X
Department of Philosophy, Ibn Haldun University, Başak Mah. Ordu Cad. F-05 Blok No:3 P.K. 34480 Başakşehir/İstanbul, Türkiye; Department of Philosophy, Middle East Technical University, Üniversiteler Mahallesi, Dumlupınar Bulvarı No:1, 06800 Çankaya/Ankara, Türkiye
selenfettahoglu@gmail.com

among scientists, we refer to natural kinds. Making use of natural kinds lies at the heart of the scientific practice for many theorists and philosophers of science. One reason for this is that natural kinds play an important role in predictions, generalizations, and explanations in science. In some sciences, such as physics and chemistry, appealing to natural kinds does not seem problematic.[1] But, in this paper, I aim to say something about natural kinds in psychology. Natural kinds in psychology are crucial in some aspects. For example, if the same mental states as natural kinds can be realized by different physical states, this can be an example of multiple realization.[2] Here, my focus will be on cognition being a natural kind.

Defining cognition is troublesome. Coarsely, by cognition I mean the mental processes that may give rise to behavioral flexibility. My understanding of cognition in terms of behavioral flexibility is based on Cameron Buckner's approach in "A property cluster theory of cognition."[3] Buckner contends that we can take cognition as a natural kind. Examples of behavioral flexibility, he argues, include the following: context-sensitivity, speed,

---

1 For example, there are different token (or particular) molecules of water, but we take H2O as a natural kind, to speak about all water molecules in the universe and track their regularities. Classifications and natural kinds are important for many scientific disciplines, which display probably the most elegant appearance in chemistry. Elements are classified in the periodic table and this classification allows us to track the regularities in the behaviors of the elements which are dispersed throughout the universe. Also, natural kinds can have useful applications in biology, such as the classification of species and their organs. In terms of organs, the classification of kidneys as natural kinds can help us monitor their regularities and differences among different people. However, the issue that I will be discussing in this article is natural kinds in psychology.

2 Multiple realization could support the functionalist theory of mind, which implies the possibility of cognition in artificial intelligence. Simply put, if the same mental state can be realized by different physical states, a silicon computer could possibly realize the same mental state as well. In other words, if cognition can be taken as a natural kind, and if it is multiply realized, then it could be realized in a silicon computer as well. This has crucial implications for philosophical discussions about the ethics of artificial intelligence and the potential risks caused by it.

3 Buckner's approach is not an exception. Many scholars consider behavioral flexibility as an indicator of cognition. Emery and Clayton (2004), Kristin Andrews (2020, 12), and Albert Newen (2015; Newen et al. 2018) are some examples.

class formation, abstract learning, multimodality, inhibition of behavior, monotonic integration, and, expectation generation and monitoring. According to Buckner's homeostatic property cluster theory (HPC)[4], being a natural kind does not require sharing every property listed here. Instead, the presence of a cluster of properties is enough to identify a natural kind. Behavioral flexibility constitutes just such a cluster, a cluster sufficient to consider animal cognition as a natural kind (see Buckner 2015).

From a physicalist point of view, one might think that we should apprehend psychological kinds as we understand biological kinds. When it comes to natural kinds in biology, the common approach is to base classification on the evolutionary tree (Khalidi 2023, 44). Biologists usually make the classification of the natural kinds of species by referring to a common ancestor.[5] I do not object to this in this article. But when it comes to psychology[6], I think using only common ancestry to uncover natural kinds is the wrong strategy. I aim to demonstrate, using Daniel Dennett's *Real Patterns* account, that cognition, which evolves convergently, is a natural kind. Therefore, cognition does not depend on ancestry for being a natural kind.

To the best of my knowledge, no other study claims that cognition is a real pattern. Thus, although the claim that cognition is a natural kind is defended elsewhere, my article is original in terms of its perspective. My approach draws upon the convergent evolution of cognition among birds and mammals.

---

[4] This view is proposed as an alternative view to relying on necessary and sufficient conditions for identifying natural kinds. Richard Boyd defends this view (see Boyd 1991).

[5] Taking biological species as a natural kind by applying to common ancestry is a plausible way to talk about them (Khalidi 2023, 41-42; see LaPorte 2009). However, I do not disregard that whether biological species are examples of natural kind is contentious. For example, Michael Ghiselin and David Hull think that species are not natural kinds but individuals (see Ghiselin 1974; Hull 1976). Still, even if species are individuals and not natural kinds, this would not harm my central thesis, which is the claim that cognition is a natural kind. I mention biological species just as an example of opposition to Cameron Buckner's claim. Without this example, I can still defend my thesis, which is that cognition is a real pattern and, therefore, is a natural kind.

[6] Although I do think that my strategy here can be extended to biology, my focus here will be the natural kinds in psychology.

2. Convergent Evolution

All living beings on Earth have a common ancestor (Akanuma 2019). Throughout evolution, living beings diversified and different functions and organs evolved. Many independent branches of the evolutionary tree produced dissimilar traits. And similar traits are usually homologous (Butler 2009). Yet there are many examples of convergent evolution. Convergent evolution refers to the evolution of similar characteristics in different branches of the evolutionary tree. Some traits evolve repeatedly in independent branches of the evolutionary tree, despite being absent in their last shared common ancestor. The reason for this is the power of natural selection. Non-functional and developmentally unsustainable traits cannot survive and because of the restricted number of functionally possible characters, we see that similar functions occur repeatedly. For example, wings of bats and birds are convergent, wings were not observable in the last common ancestor they shared (Seed et. al. 2009). Likewise, legs evolved in arthropods and vertebrates through convergent evolution (Ritzmann et. al. 2004). Also, the camera-eye evolved convergently in mammals and octopi, although their last common ancestor lacked a camera eye (Clayton and Emery 2008, 130).

In this article, I will focus on the convergent evolution of cognition among birds and mammals. Birds have very high-level cognitive capacities. I should note that the ideal comparison would involve two single species but, to the best of my knowledge, there is no scientific comparison of two species from the aspects that I am interested in. So, to narrow down the compared examples, I will compare corvids and apes. I will argue that corvids and apes share the same natural kind of cognition.

To begin, we will start by exemplifying some impressive abilities of corvids and apes. The manufacture and use of tools, which require the integration of visual and haptic information, are seen as a cognitive capacity (Buckner 2015, 320). Among apes, chimpanzees display behaviors of tool usage such as ant-dipping and termite fishing (McGrew 1992). Likewise, New Caledonian crows among corvids, excel at tool usage and manufacturing (Hunt 1996). Crows and ravens can also delay gratification for a quality of reward (Hillemann et al. 2014), which indicates impulse control through inhibitory behavior. Similarly, chimpanzees can delay gratification for several minutes for a quantity of reward (Dufour et. al. 2007). Moreover, both corvids and chimpanzees display transitive inference (see Paz-y-Mino C et al. 2004; Gillan 1981). These behaviors are thought of as cognitive, as psychologists take these as indicators of behavioral flexibility (Buckner 2015). Studies on corvids suggest that they have beyond instinctive behaviors, with very high cognitive capacities. For example, scrub-jays seem to have the theory of mind, an efficient memory, can mentally travel in time, and plan for the future (Baciadonna et. al. 2021). The cognitive capacities of corvids are comparable to that of apes (Güntürkün 2012).

It is easier to conceive of apes as cognitive beings because they separated from us relatively recently, from 9 million to 5 million years ago (Andrews 2019). However, the last common ancestor of birds and mammals lived around 300 million years ago, and that ancestor is not considered cognitive (Emery 2016, 38). Thus, cognition evolved separately in different branches of the evolutionary tree. Here, we have a case of convergent evolution (Van Horik et al. 2012).

At this point, consider a possible objection: one could claim that since cognition evolved slowly, some degree of cognition was already present in the common ancestor of corvids and apes; and, any claim about the common ancestor's lack of cognition is troublesome. At first sight, this would seem to hinder my claim that the natural kind of cognition evolved convergently in corvids and apes. However, because of the definition of cognition I adopted here, I think the slow evolution of cognition does not harm my thesis. Following Buckner (2015), I define cognition as a cluster of properties. As long as any cluster of properties among the examples of behavioral flexibility is satisfied, we have cognition. Although some degree of behavioral flexibility may be present in the common ancestor of corvids and apes, if the properties of behavioral flexibility do not occur as a cluster, we can say that there is no cognition. Buckner's property cluster theory of cognition provides a threshold for the presence of cognition[7], and it is

[7] This is similar to autism, which appears as a degree, being a natural kind (see Khalidi 2023, 60).

unlikely that the common ancestor of corvids and apes had the cluster of properties for complex cognition (see Emery and Clayton 2004)[8].

Scientists believe that cognition among birds evolved convergently primarily because cognitive capacities in mammals occur in the prefrontal cortex, yet birds do not have a prefrontal cortex. Instead of the prefrontal cortex, birds evolved another region (nidopallium caudolaterale) for cognitive functions. It is believed that this region evolved separately from the mammalian prefrontal cortex as it has a structure different from the prefrontal cortex (Clayton and Emery 2008, 131-132; Güntürkün et al. 2024; Güntürkün 2012). Moreover, some corvids have more complex cognitive capacities than some mammals (Clayton and Emery 2008, 130), and this, too, implies convergent evolution.

The similarity of cognitive capacities between corvids and apes seems to suggest multiple realization. Yet differences may imply cognition is not a natural kind. As mentioned before, Cameron Buckner (2015) claims that cognition displays a cluster of properties (some of which are exemplified above), which suggest that cognition in corvids and apes is a natural kind. However, because they result from different evolutionay paths, one might object to cognition in corvids and apes being the same natural kind. Since biologists classified natural kinds according to common ancestry, it is not unreasonable to expect the application of the same strategy to psychology as well. Buckner's claim is not applicable to convergently evolved cognition if natural kinds are identified by common ancestry. Indeed, some have proposed that pain in dogs and pain in humans are of different natural kinds because they are in different species (see Kim 1972, 190; Lewis 1969). As such, cognition in birds and in apes must be different.

Now, I will employ Dennett's *Real Patterns* account to support the claim that although these two sorts of cognitions follow different evolutionary paths through convergent evolution, they are, nonetheless, of the same natural kind.

[8] Here, I think what is cognitive in Emery and Clayton's article is also cognitive in Buckner's theory. The criteria in these two articles are compatible. They both emphasize behavioral flexibility in their understanding of cognition.

3. Cognition as a Real Pattern?

In 1991, Daniel Dennett wrote "Real Patterns", which states that compressible patterns that allow prediction are real. Dennett considers whether beliefs are real and because belief attribution can help us to explain complex behavioral patterns in a compressed way, he thinks they are. Being compressible means that instead of counting one by one, we can express the data in a shorter description. A pattern that is nonrandom and allows for prediction is real. Since beliefs allow the prediction of behaviors, they are real.

I will now argue that cognition is real in the same way. Although Dennett's article has been criticized because it does not offer sufficient conditions for being a real pattern (Ladyman and Ross 2007, 205), I will endorse his predictability and compressibility approach to claim the initial plausibility that cognition is a real pattern. My aim is to propose that cognition is a real pattern, and to underline the plausibility that convergent evolution can produce real patterns that may have implications for natural kinds.

Ladyman and Ross (2007), with more detailed descriptions of patternhood,[9] claim that natural kinds are real patterns: "We contend that everything a naturalist could legitimately want from the concept of a natural kind can be had simply by real patterns" (Ladyman and Ross 2007, 294). They define natural kinds as "real patterns with a high indexical redundancy" (Ladyman and Ross 2007, 295). A pattern has low indexical redundancy if it can be observed at just one point in space and time (Cretu 2020, 4). In that case, convergence of cognition seems to increase its indexical redundancy because repetition increases the number of measurements. Cretu's definition of real patternhood gives us a clear view: "Real patterns are robust relations amongst entities exhibited by any two given entities with sufficient regularity at any given scale" (Cretu 2020, 21). The main question is whether cognition is a real pattern. Although not conclusive, I offer reasons to think it is.

---

9 They take projectability as a condition for being a real pattern. Projectability refers to the possibility of calculating an outcome from something else. As I understand it, it could be used interchangeably with predictability here. For example, the instances of clustered properties are calculatable from the reality of cognition. Still, I am not claiming that my approach here satisfies their requirement for being a real pattern.

Does cognition have predictive power? According to Buckner's property cluster theory of cognition, cognition as a natural kind displays some behavioral characteristics that cluster. Although not all of them need to be displayed in every cognitive animal, a cluster of the properties should be displayed. As I mentioned above, corvids and apes seem to display some of the properties. Buckner signifies; context-sensitive behaviors, rapidity of learning, categorization capacity, abstract learning, multimodality, inhibition of behavior, monotonic integration (such as transitive inference) and, expectation generation and monitoring, as clustered properties of cognition. A cognitive animal must have some but not all of these properties (Buckner 2015).

If, because clustered, we take cognition as a natural kind, we can predict behaviors (just as we predict behaviors based on beliefs). For example, if we have reasons to think that an animal has cognition, we would expect it to display some instances of behavioral flexibility.[10] For example, we can expect a cognitive animal to develop a strategy to hinder pilfering of its food. If beliefs are real as Dennett stated, then, I claim, cognition can be real as well. If so, cognition is a natural kind. We should be able to take convergently evolved examples of cognition as the same natural kind because they display real patternhood.

4. Some Other Words About Prediction

Dennett offers *The Game of Life* as an example of a pattern generation process. *The Game of Life* is a computer simulation that proceeds with a simple rule through which patterns are generated, and new shapes emerge: "There are computer simulations of the Life world in which one can set up configurations on the screen and then watch them evolve according to the single rule" (Dennett 1991, 38). The Game of Life, which is intended to

[10] Of course, we infer that an animal has cognition *because* it displays behavioral flexibility. So, expecting behavioral flexibility because an animal is cognitive does not seem a novel prediction. However, what I mean here is that if we infer that an animal has behavioral flexibility, we can predict how it will behave in the contexts that we have never tested before. So, this is a novel prediction.

represent evolution, gives predictions for the upcoming figure, thanks to the rule it applies. In evolution, the primary pattern generation process is natural selection. The convergent evolution of many organs and cognition likewise shows us the power of natural selection. The many examples of convergent evolution can even be used to predict the evolution of possible extraterrestrials.[11] With sufficient information about their environments, it could be possible to predict the body shapes of extraterrestrial beings and even their cognition. If natural selection produces similar outcomes repeatedly on Earth, the same rule would produce similar outcomes on an extraterrestrial planet as well, if the environment is similar (see Morris 2003; Kershenbaum 2021).

The multiplicity of examples of convergent evolution in the biological realm gives rise to ideas about the predictability of evolution. Natural kinds can help us with predictability. For example, the periodic table of elements in chemistry allows us to make predictions. Likewise, convergent evolution allows biologists to make a "periodic table" of life (see McGhee 2008; McGhee 2011, 261). As the periodic table of elements can be used to predict the behavior of elements, a periodic table of life can be used to predict the direction of evolution (McGhee 2011, 276).[12] In the case of cognition, by acknowledging cognition as a natural kind, we can track the regularity of its presence in the evolutionary process. For example, as mentioned above, we can predict the appearance of cognition on an extraterrestrial planet (see Morris 2003; Kershenbaum 2021; Schulze-Makuch and Bains 2017; Powell 2020).

Another question about the real patternhood of cognition is compressibility (Dennett 1991). I will not delve into the issue of compressibility in detail, but I will say a few words about its initial plausibility in cognition. Compressibility is possible if the data are not random. It means that to be real, the pattern has to have low entropy. Entropy can be considered a measure of disorder. And order can be described as "a constant relation

[11] This is a contentious claim, but I do not think it is wrong to exemplify this as I do not claim a conclusive position in this article.

[12] This will not give us %100 predictions. But Daniel Dennett's account allows real patterns with less than %100 predictions.

between neighboring constituent elements of a system" (Grandpierre 2005). If there is no constant relation, a system is random.

Is there a relation among the cognitive behaviors that are clustered? Because of the claim that they are clustered (see Buckner 2015), a cognitive system has low entropy. If we compare the behaviors caused by cognition in corvids and apes with the behaviors caused by non-cognitive processes, such as the behaviors of bacteria, we can see that there is a non-random relation in the behaviors caused by cognition. They occur together. I will not offer any quantitative calculations, as just exemplifying the common qualities displayed by cognitive capacities seems enough to support the compressibility of cognition. Still, I do not claim a conclusive position, as my aim here is to show the initial plausibility of cognition being a real pattern, and if a real pattern, a natural kind.

5. Conclusion

I have proposed a way, following Buckner's property cluster theory of cognition, to take convergently evolved examples of cognition as a natural kind. However, in biology, natural kinds are usually classified according to common ancestry. Therefore, whether convergently evolved examples of cognition are of the same natural kind is debatable. I proposed that Ladyman and Ross's method of taking natural kinds as real patterns could ease taking cognition as a natural kind. To this end, in the first section, I introduced the importance of natural kinds. In the second section, I explained what convergent evolution is, and gave an example case, corvids and apes, to demonstrate that thanks to convergent evolution similar cognitive capacities evolved. Then, I mentioned the idea that cognition should be taken as a natural kind. In the last two sections, I proposed the idea that cognition could be an example of a real pattern. My discussion about the real patternhood of cognition is far from conclusive and should be understood more as a proposal than a conclusion. If further studies support the idea that cognition is a real pattern, this would support the view that it is a natural kind, the same natural kind in corvids and apes, although it evolved convergently in separate branches of evolution.

Acknowledgements

I am grateful to Gerardo Viera, whose help was significant for the publication of this article. I am thankful for his guidance and revisions. Likewise, Kelly James Clark meticulously revised and gave invaluable feedback to this article. I deeply appreciate his support. Also, I thank my current advisor Ayhan Sol and former advisor Enis Doko for their help and feedback in this article. I also appreciate the feedback from one anonymous referee.

References

Andrews, Kristin. 2020. *The Animal Mind: An Introduction to the Philosophy of Animal Cognition*. Second edition. New York: Routledge.

Akanuma, Satoshi. 2019. "The Common Ancestor of All Modern Life." In *Astrobiology*, 91-103. Springer Singapore. https://doi.org/10.1007/978-981-13-3639-3_7

Andrews, Peter. 2019. "Last Common Ancestor of Apes and Humans: Morphology and Environment." *Folia Primatologica*, 91(2): 122–48. https://doi.org/10.1159/000501557

Baciadonna, Luigi, Francesca M. Cornero, Nathan J. Emery, and Nicola S. Clayton. 2021. "Convergent Evolution of Complex Cognition: Insights from the Field of Avian Cognition into the Study of Self-Awareness." *Learning & Behavior* 49(1): 9–22. https://doi.org/10.3758/s13420-020-00434-5

Buckner, Cameron. 2015. "A Property Cluster Theory of Cognition." *Philosophical Psychology* 28 (3): 307–36. https://doi.org/10.1080/09515089.2013.843274

Butler, Ann B. 2009. "Evolution and the Concept of Homology." In *Encyclopedia of Neuroscience*, edited by Marc D. Binder, Nobutaka Hirokawa, and Uwe Windhorst, 1208-1212. Berlin and Heidelberg: Springer. https://doi.org/10.1007/978-3-540-29678-2_3117

Boyd, Richard. 1991. "Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds." *Philosophical Studies*, 61 (1-2): 127–48. https://doi.org/10.1007/BF00385837

Clayton, Nicola, and Nathan Emery. 2008. "Canny Corvids and Political Primates." In *The Deep Structure Of Biology,* edited by Simon Conway Morris, 17–31. West Conshohocken, Pennsylvania: Templeton Foundation Press.

Cretu, Ana-Maria. 2020. "Natural Kinds as Real Patterns: Or How to Solve the Commitment Problem for Perspectival Realism." Preprint, submitted January 18, 2020. http://philsci-archive.pitt.edu/16828/

Dennett, Daniel. 1991. "Real Patterns." *The Journal of Philosophy*, 88 (January): 27–51.

Dufour, V., M. Pelé, E. H. M. Sterck, and B. Thierry. 2007. "Chimpanzee (Pan Troglodytes) Anticipation of Food Return: Coping with Waiting Time in an Exchange Task." *Journal of Comparative Psychology*, 121: 145–55. https://doi.org/10.1037/0735-7036.121.2.145

Emery, Nathan J., and Nicola S. Clayton. 2004. "The Mentality of Crows: Convergent Evolution of Intelligence in Corvids and Apes." *Science*, 306 (5703): 1903–7. https://doi.org/10.1126/science.1098410

Emery, Nathan. 2016. *Bird Brain: An Exploration of Avian Intelligence.* Princeton, New Jersey: Princeton University Press.

Ghiselin, Michael T. 1974. "A Radical Solution to the Species Problem." *Systematic Zoology*, 23: 536–44.

Gillan, Douglas J. 1981. "Reasoning in the Chimpanzee: II. Transitive Inference." *Journal of Experimental Psychology: Animal Behavior Processes*, 7: 150–64. https://doi.org/10.1037/0097-7403.7.2.150

Grandpierre, Attila. 2005. "Order, Structure, Physical Organization and Biological Organization." *Frontiers Perspectives*. Available at https://old.konkoly.hu/staff/grandpierre/Order.pdf

Güntürkün, Onur. 2012. "The Convergent Evolution of Neural Substrates for Cognition." *Psychological Research*, 76 (2): 212-219. https://doi.org/10.1007/s00426-011-0377-9

Güntürkün, Onur, Roland Pusch, and Jonas Rose. 2024. "Why Birds Are Smart." *Trends in Cognitive Sciences*, 28 (3): 197–209. https://doi.org/10.1016/j.tics.2023.11.002

Hillemann, Friederike, Thomas Bugnyar, Kurt Kotrschal, and Claudia A. F. Wascher. 2014. "Waiting for Better, Not for More: Corvids Respond to Quality in Two Delay Maintenance Tasks." *Animal Behaviour*, 90 (April): 1–10. https://doi.org/10.1016/j.anbehav.2014.01.007

Hull, David L. 1976. "Are Species Really Individuals?" *Systematic Zoology*, 25 (2): 174. https://doi.org/10.2307/2412744

Hunt, Gavin R. 1996. "Manufacture and Use of Hook-Tools by New Caledonian Crows." *Nature*, 379 (6562): 249–51. https://doi.org/10.1038/379249a0

Kershenbaum, Arik. 2021. *The Zoologist's Guide to the Galaxy: What Animals on Earth Reveal About Aliens--and Ourselves.* London: Penguin Publishing Group.

Khalidi, Muhammad Ali. 2023. *Natural Kinds*. 1st ed. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009008655

Kim, Jaegwon. 1972. "Phenomenal Properties, Psychophysical Laws, and the Identity Theory," edited by Sherwood J. B. Sugden. *Monist*, 56 (2): 177–92. https://doi.org/10.5840/monist19725623

Ladyman, James, Don Ross, David Spurrett, and John G. Collier. 2007. *Every Thing Must Go: Metaphysics Naturalized*. Oxford and New York: Oxford University Press.

Laporte, Joseph. 2009. *Natural Kinds and Conceptual Change*. Cambridge: Cambridge University Press.

Lewis, David. 1969. "Review of *Art, Mind, and Religion*." *Journal of Philosophy*, 66 (1): 22–7. https://doi.org/10.2307/2024154

McGhee, George. 2008. "Convergent Evolution: A Periodic Table of Life." In *The Deep Structure Of Biology,* edited by Simon Conway Morris, 17–31. Pennsylvania: Templeton Foundation Press.

McGhee, George R. 2011. *Convergent Evolution: Limited Forms Most Beautiful*. Cambridge, Massachusetts: MIT Press.

McGrew, William Clement. 1992. *Chimpanzee Material Culture: Implications for Human Evolution*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511565519

Morris, Simon Conway. 2003. *Life's Solution: Inevitable Humans in a Lonely Universe*. Cambridge: Cambridge University Press.

Newen, Albert. 2015. "What Are Cognitive Processes? An Example-Based Approach." *Synthese*, 194 (11): 4251–68. https://doi.org/10.1007/s11229-015-0812-3

Newen, Albert, Shaun Gallagher, and Leon De Bruin. 2018. "4E Cognition: Historical Roots, Key Concepts, and Central Issues." In *The Oxford Handbook of 4E Cognition*, edited by Albert Newen, Leon De Bruin, and Shaun Gallagher, 2–16. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198735410.013.1

Paz-y-Miño C, Guillermo, Alan B. Bond, Alan C. Kamil, and Russell P. Balda. 2004. "Pinyon Jays Use Transitive Inference to Predict Social Dominance." *Nature*, 430 (7001): 778–81. https://doi.org/10.1038/nature02723

Powell, Russell. 2020. *Contingency and Convergence: Toward a Cosmic Biology of Body and Mind*. Vienna Series in Theoretical Biology. Cambridge, Massachusetts: The MIT Press.

Ritzmann, Roy E., Roger D. Quinn, and Martin S. Fischer. 2004. "Convergent Evolution and Locomotion through Complex Terrain by Insects, Vertebrates and Robots." *Arthropod Structure & Development*, 33 (3): 361–79. https://doi.org/10.1016/j.asd.2004.05.001

Schulze-Makuch, Dirk, and William Bains. 2017. *The Cosmic Zoo*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-62045-9

Seed, Amanda, Nathan Emery, and Nicola Clayton. 2009. "Intelligence in Corvids and Apes: A Case of Convergent Evolution?" *Ethology*, 115 (5): 401–20. https://doi.org/10.1111/j.1439-0310.2009.01644.x

Van Horik, Jayden O., Nicola S. Clayton, and Nathan J. Emery. 2012. "Convergent Evolution of Cognition in Corvids, Apes and Other Animals." In *The Oxford Handbook of Comparative Evolutionary Psychology*, edited by Todd K. Shackelford and Jennifer Vonk, 1st ed., 80–101. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199738182.013.0005

Organon F 31 (2) 2024: 197–201
https://doi.org/10.31577/orgf.2024.31205
ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

DISCUSSION NOTE

No Path from Modal Rationalism to Fundamental Scrutability

Darren Bradley*

Received: 10 August 2023 / Accepted: 15 April 2024

Abstract: Gabriel Rabin (2020) offers an argument from Modal Rationalism to Fundamental Scrutability. I show that the argument is invalid as stated. I offer two ways of strengthening the argument but argue that neither is effective.

Keywords: Modal rationalism; scrutability, fundamentality; B-type physicalism.

Gabriel Rabin (2020) offers the following argument:

1. *A Priori Access/Modal rationalism*:
An idealized reasoner could, in principle, completely describe each and every possible world down to the finest detail

2. *Supervenience on the Fundamental*:
No two worlds differ without differing at the fundamental level

Therefore

* University of Leeds
https://orcid.org/0000-0002-7757-3902
School of Philosophy, Religion and History of Science, University of Leeds, Leeds LS2 9JT, United Kingdom
D.J.Bradley@leeds.ac.uk

3. *Fundamental Scrutability*:

For each world w, a conditional Fw→Tw is knowable a priori, where Fw is a complete description of w's fundamental level and Tw is the set of all truths at w.

I think this argument is invalid. To start with a diagnosis in abstract terms, *Modal Rationalism* is about epistemic possibility while *Supervenience on the Fundamental* is about metaphysical possibility. They do not allow us to derive *Fundamental Scrutability* which connects epistemic and metaphysical possibility.

Rabin explains why he thinks *Fundamental Scrutability* follows from the premises as follows:

> We give [idealized reasoner Athena] Fw, and ask her to a priori reason her way to Tw. Here's how she can do so. By *A Priori Access*, she can describe all the ways the possible worlds could be. Therefore, she knows that there is a world, call it v, at which Fw & Tw. But to deduce Tw from Fw she needs to figure out that v is the only world at which Fw. Might it be that Fw & Tx, for some x =/= w? Absolutely not, by *Supervenience on the Fundamental*. Fw & Tx is impossible.

But *Supervenience on the Fundamental* is not enough to arrive at this negative answer. The easiest way to see why is to consider what Chalmers (2003 section 5) calls type-B physicalists, who affirm that (physical) fundamental properties *metaphysically* necessitate non-fundamental phenomenal properties (supervenience), but do not *epistemically* necessitate these non-fundamental phenomenal properties (scrutability).[1] To use a familiar example, suppose the firing of c-fibres metaphysically necessitates pain, but agents cannot infer a priori from the firing of c-fibres to the instantiation of pain.

So, let's re-write the passage above using the pain/c-fibres example:

> We give Athena a physical description of the world, and ask her to a priori reason her way to a phenomenal description of the world. Here's how she can do so. By *A Priori Access*, she can describe all the ways the possible

[1] See Chalmers (2012) for this terminology.

> worlds could be. Therefore, she knows that there is a world, call it v, at which c-fibres fire & pain is instantiated. But to deduce the pain from the c-fibres, she needs to figure out that v is the only world with that physical description. Might it be that c-fibres fire without pain?

The B-type phisicalist says yes! Athena can work out *that* there is only one metaphysically possible world fitting the physical description, but she cannot work out *whether* it contains pain. Both a world with pain and without pain are epistemically possible, and as we are asking what an agent can infer, it is epistemic possibility which matters. (The type-B physicalist will agree that Fw & Tx is *metaphysically* impossible, but *Fundamental Scrutability* requires that it is epistemically impossible.)

What would be needed to make the argument valid? I'll consider two ways to strengthen the argument, but neither will be very effective at supporting *Fundamental Scrutability.*

First, we could strengthen *Supervenience on the Fundamental:*

> *Supervenience on the Fundamental+*:
> No two *epistemically possible* worlds differ without differing at the fundamental level

This would close the gap between epistemic and metaphysical possibility, saying that when we fix the fundamental level, we fix which world is epistemically possibly at all levels.

But no-one who doubts *Fundamental Scrutability* will find *Supervenience on the Fundamental+* tempting. Type-B physcialists who deny that there is an a priori path of reasoning from the physical to the mental are saying that there *are* two different epistemically possible worlds which don't differ at the fundamental level.

Furthermore, it would be misleading to call *Supervenience on the Fundamental+* a 'supervenience' principle, as supervenience has usually been connected to metaphysical possibility e.g. Mooreans agree that the ethical supervenes on the physical but deny that the ethical could be explained by the physical, leading to the open question argument and non-naturalism.[2]

[2] See Enoch (2011) for a contemporary defence of non-naturalism.

Alternatively, *Modal Rationalism* could be strengthened:

A Priori Access/Modal rationalism+:
An idealized reasoner could, in principle, identify whether a world were metaphysically possible given a complete description of it in canonical[3] terminology.

This would also close the gap between epistemic and metaphysical possibility, allowing an idealized reasoner to tell whether 'c-fibres firing without pain' is metaphysically possible. I suspect that this is the way modal rationalists will want to go, as a motivation behind modal rationalism is that we have epistemic access to all the facts about modality.

But type-B physicalists will deny that an idealized reasoner could, in principle, identify whether a world with c-fibres firing and without pain is metaphysically possible. They hold that such a world is metaphysically impossible, but that one cannot discover this a priori. There are other examples not related to the mind-body problem. Perhaps some controversial metaphysical theses (e.g. the existence of numbers, the existence of God, principles of composition) are metaphysically necessary but epistemically contingent. If so, an idealized reasoner might be unable to identify whether a world without numbers is metaphysically possible despite having a complete description of the world in canonical vocabulary. Thus, an idealized reasoner's list of all the epistemically possible worlds in canonical vocabulary would include some which are metaphysically impossible.

To be clear, I am actually sympathetic to *Fundamental Scrutability.* My point here is that it does not follow from Rabin's versions of *Modal Rationalism* and *Supervenience on the Fundamental.* We need stronger premises to rule out type-B physicalists who insist that epistemic possibilities outrun metaphysical possibilities; but the natural stronger formulations discussed above will not be tempting to opponents of *Fundamental Scrutability.*

---

3 Without this restriction to canonical terminology this principle would be trivially true. Plausible canonical terminology would be the conjunction of physical, phenomenal and indexical terminology plus a that's-all clause. See Chalmers (2012 section 8.5) for a detailed discussion of various version of modal rationalism.

Ackowledment

I am grateful to the Templeton Foundation for partially funding this research.

References

Chalmers, D. (2003). 'Consciousness and Its Place in Nature'. In *Blackwell Guide to the Philosophy of Mind*, edited by S. Stich and F. Warfield, 102–42. Oxford: Blackwell.

Chalmers, D.J. (2012). *Constructing the World*. Oxford: Oxford University Press.

Enoch, D. (2011). *Taking Morality Seriously: A Defense of Robust Realism*. Oxford: Oxford University Press on Demand.

Rabin, G. O. (2020). A Short Argument from Modal Rationalism to Fundamental Scrutability. *Thought: A Journal of Philosophy*, 9(2), 137–39. https://doi.org/10.1002/tht3.446