METHODOLOGICAL NATURALISM AND REFLEXIVITY REQUIREMENT

Hamed BIKARAAN-BEHESHT

ABSTRACT: Methodological naturalists regard scientific method as the only effective way of acquiring knowledge. Quite the contrary, traditional analytic philosophers reject employing scientific method in philosophy as illegitimate unless it is justified by the traditional methods. One of their attacks on methodological naturalism is the objection that it is either incoherent or viciously circular: any argument that may be offered for methodological naturalism either employs a priori methods or involves a vicious circle that ensues from employing the very method that the argument is aimed to show its credentials. The charge of circularity has also been brought against the naturalistic arguments for specific scientific methods; like the inductive argument for induction and the abductive argument for the inference to the best explanation. In this paper, I respond to the charge of circularity using a meta-methodological rule that I call 'reflexivity requirement.' Giving two examples of philosophical works, I illustrate how the requirement has already been considered to be necessary for self-referential theories. At the end, I put forward a meta-philosophical explanation of the naturalism-traditionalism debate over the legitimate method of philosophy.

KEYWORDS: methodological naturalism, reflexivity, scientific method, armchair philosophy, a posteriori argument

1. Introduction: Methodological Naturalism

Methodological naturalism is frequently defined as a commitment to employ scientific method – or more generally, empirical methods – in philosophy as well as science. It means that philosophers, like scientists, should exclusively make use of scientific method to address philosophical problems, since from the naturalists' point of view, it is the only effective way of acquiring knowledge of the world. Such a reading of methodological naturalism has already appeared in many works concerning naturalism. For example, in *Encyclopedia of Philosophy*, Keith Campbell views the methodological aspect of naturalism in this way:

> […] naturalism is sometimes regarded as a rule of method rather than a metaphysical doctrine. There is a natural method of inquiry, which consists in setting out to explain and understand the world by finding the natural causal processes by which natural objects come into being, produce their effects, and pass

> away. *All genuine knowledge* is of this natural, experimental kind; […].[1]

Jaegwon Kim also defines 'the epistemological thesis' of naturalism as the idea that "*all knowledge that we can acquire* is acquirable *only* through the application of scientific method."[2] As another instance, Mario De Caro and David Macarthur characterize a "methodological (or epistemological) scientific naturalist" as someone who believes "it is *only* by following the methods of the natural sciences – or, at a minimum, the empirical methods of a posteriori inquiry – that one arrives at genuine knowledge."[3] Likewise, Michael Devitt defines naturalism as the thesis that "there is *only one way* of knowing, the empirical way that is the basis of science (whatever that way may be)."[4] And many anti-naturalists understand methodological naturalism in the same way.[5] So, I presume the following thesis to be the main claim of methodological naturalists:

> (**MN**): Scientific method is the only route to knowledge (in all areas including philosophy).

This line of thought constitutes an integral part of *philosophical* naturalism in general.[6] But, what is the reason behind the naturalists' strong belief in the view so

---

[1] Keith Campbell, "Naturalism," *Encyclopedia Of Philosophy*, Vol. 6, ed. Donald M. Borchert (2005), 492 (my italics).

[2] Jaegwon Kim, "The American Origins of Philosophical Naturalism," *Journal of Philosophical Research* 28, Supplement (2003): 83-98, 88 (my italics).

[3] Mario De Caro and David Macarthur, "Introduction: The Nature of Naturalism," in *Naturalism in Question*, ed. Mario De Caro and David Macarthur (Cambridge: Harvard University Press, 2004), 1-17, 7 (original italics).

[4] Michael Devitt, *Coming to Our Senses: A Naturalistic Program for Semantic Localism* (Cambridge: Cambridge University Press, 1996), 2; Michael Devitt, "Naturalism and the a Priori," *Philosophical Studies* 92 (1998): 45-65, 45 (my italics). See also Devitt, "Naturalism and the a Priori," 46-47.

[5] See, e.g., Robert Audi, "Philosophical Naturalism at the Turn of the Century," *Journal of Philosophical Research* 25 (2000): 27-45, 41; Laurence BonJour, "Against Materialism," in *The Waning of Materialism*, ed. Robert C. Koons and George Bealer (Oxford: Oxford University Press, 2010), 3-23, 7. For some slightly different (less extreme) definitions of methodological naturalism, see Kim Sterelny, *The Representational Theory of Mind: An Introduction* (Oxford: Basil Blackwell, 1990), xi; Tim Lewens, "A Surfeit of Naturalism," *Metaphilosophy* 43, 1-2 (2012): 46-57; De Caro and Macarthur, "Introduction;" Philip Kitcher, "Giving Darwin His Due," in *The Cambridge Companion to Darwin*, ed. Jonathan Hodge and Gregory Radick (Cambridge: Cambridge University Press, 2009), 455-476.

[6] Wilmon Henry Sheldon, "Critique of Naturalism," *The Journal of Philosophy* 42, 10 (1945): 253-270; Alexander Rosenberg, "A Field Guide to Recent Species of Naturalism," *The British Journal for the Philosophy of Science* 47, 1 (1996): 1-29; Alexander Rosenberg, "Disenchanted Naturalism," in *Contemporary Philosophical Naturalism and Its Implications*, ed. Bana Bashour and Hans D.

central to their position? That is, what is their justification for (MN)? Finding the answer to this question will become more important when it is realized that many works in philosophy draw heavily on the thesis, and there is a great controversy over it.

I start finding an answer to the question with scrutinizing (MN). (MN) has an implicit negative component: it implies that traditional methods of *armchair philosophy* or *first philosophy* fall short of directing us toward knowledge, and, as a result, they should be abandoned.[7] The main reason naturalists have for this claim is not that traditional methods are illegitimate – that is, resulting in false beliefs[8] – but it is the supposed unfruitfulness and unproductiveness of those methods.[9] Philosophy, they claim, has been extremely unsuccessful and it has been due to the infertile methods of first philosophy. On the other hand, for them, most (or even all) the knowledge we have about the world has been acquired using scientific method, that is, science achieved its success via its method.[10] Then, scientific method can be vindicated according to its past successes. If one regards such a reason as an argument for methodological naturalism – as many naturalists claim[11] – it will be an *a posteriori argument*. It is because in that argument, some evidence from history of science and history of philosophy is provided for defending (MN).

(MN) has come under two criticisms. Firstly, it is not clear what is meant by 'scientific method.' Actually, it is hard to determine a necessary and sufficient condition a kind of method should meet for being scientific. I am not going to tackle with this problem, and it does not affect what I am arguing for here. The only thing I need here is that the a posteriori way of justifying (MN) can be considered to be scientific in a sense. Secondly, contrary to what is claimed by naturalists, many anti-

---

Muller (New York: Routledge, 2014), 17-36; Kim, "American Origins;" Audi, "Philosophical Naturalism at the Turn of the Century"; Timothy Williamson, "What Is Naturalism?," in *Philosophical Methodology: The Armchair or the Laboratory?*, ed. Matthew C. Haug (New York: Routledge, 2014), 29-31; Daniel Stoljar, *Physicalism* (London: Routledge, 2010), 11.

[7] Willard Van Orman Quine, *Theories and Things* (Cambridge, MA: Harvard University Press, 1981), 72.

[8] It seems that some naturalists even tend to make that claim. See, e.g., Daniel C. Dennett, "Current Issues in the Philosophy of Mind," *American Philosophical Quarterly* 15, 4 (1978): 249-261, 250.

[9] See John Dewey, *The Influence of Darwin on Philosophy and Other Essays in Contemporary Thought* (New York: Henry Holt and Company, 1910), 17-18.

[10] See, e.g., Alexander Rosenberg, "Can Naturalism Save the Humanities?," in *Philosophical Methodology: The Armchair or the Laboratory?*, ed. Matthew C Haug (New York: Routledge, 2014), 39-42.

[11] See Ronald N. Giere, "Modest Evolutionary Naturalism," *Biological Theory* 1, 1 (2006): 52-60, 53; Hilary Kornblith, "Naturalism: Both Metaphysical and Epistemological," *Midwest Studies in Philosophy* 19, 1 (1994): 39-52, 49.

naturalists believe that the method of philosophy must be certain and scientific method whose infallibility is not beyond doubt is not legitimate in philosophy. (I will call these anti-naturalists *methodological traditionalists* – traditionalists for short.[12]) Hence, not only do traditionalists want to undermine (MN), but also they insist on a more extreme view about the illegitimacy of using scientific method in philosophy such that they even reject the more moderate versions of naturalism. However, one of the most important counterarguments of traditionalists against methodological naturalism has been posed against (MN). As explained in the next section, via that argument, traditionalists aim to block *any* kind of argument naturalists *may* adduce for methodological naturalism. If successful, such an argument deprives naturalists of having any argument for their position.

2. *A Priori Argument* against Methodological Naturalism: The Charge of Circularity

The simple form of the counterargument against methodological naturalism proceeds with a dilemma: naturalists either justify (MN) *a priori* or justify it *a posteriori*. According to (MN) itself, a priori justification is not allowed. Then, the first horn of the dilemma leads naturalists to some kind of incoherence.[13] Therefore, naturalists have no choice rather than the second horn, i.e. to go on an a posteriori argument. As described in the previous section, this is their actual choice. But, is a posteriori justification legitimate in philosophy? As said before, according to *standard* armchair philosophy, the answer is negative unless there is some a priori argument for legitimacy of a posteriori justification. Now, if naturalists say that they consider a posteriori justifications to be as legitimate as a priori ones, then,

[12] Siegel ("Naturalized Epistemology and 'First Philosophy'," *Metaphilosophy* 26, 1-2 (1995): 46-62, esp. 48-49; "Empirical Psychology, Naturalized Epistemology, and First Philosophy," *Philosophy of Science* 51, 4 (1984): 667-676, esp. 671-673) rejects such a characterization of traditionalism. However, even if he is right, his idea of 'moderate first philosophy' – including such ideas as 'justification from without' and 'nonempirical theory of justification' – may be used here to characterize traditionalism without having a significant effect on my discussion. It is because any idea that seeks an 'extrascientific basis' for science will not be acceptable from a naturalistic point of view (see Giere, "Modest Evolutionary," 53), since naturalists are in conflict with traditionalists over "the (il)legitimacy of appealing to the results of scientific inquiry in order to establish the epistemic credentials of (the results of) that selfsame inquiry" (Siegel, "Naturalized Epistemology," 49 (fn. 45)). Then, whichever way we understand traditionalism (or first philosophy), it is no less than a struggle to analyze science 'from without' to avoid "a seemingly vicious circularity problem" which naturalists deny (see Siegel, "Naturalized Epistemology," 52. See also Paul A. Roth, "Siegel on Naturalized Epistemology and Natural Science," *Philosophy of Science* 50, 3 (1983): 482-493).

[13] See Giere, "Modest Evolutionary," 53.

traditionalists claim, they indulge themselves in a vicious circle, because they suppose the claim they are supposed to prove. So, in either horn, naturalists are unable to justify their position, and then, they are unjustified. The situation can be called a 'naturalist's dilemma.'[14]

The dilemma has been touched on by some anti-naturalists; among them are Robert Audi and Harvey Siegel. Audi says that if we characterize naturalism as prescribing 'certain basic methods of inquiry,' then, compared with other methods, it considers those methods privileged and as a result, naturalism is subject to 'actual-method chauvinism' – the claim that the only legitimate methods are some specific actual methods of science. Then, he continues:

> I find it hard to see how to justify such privilege except possibly on a priori grounds. […] unless we are to acquiesce in skepticism, we must apparently choose between *a priori justification* on the one hand and, on the other, either forgoing justification or countenancing *circular justification*.[15]

The dilemma Audi describes is exactly the naturalist's dilemma.

Siegel, when criticizing Gibson's defense of Quine's naturalism, makes use of the same tactic to disarm naturalists: "if the premises of any such [naturalistic] argument [for naturalism] are justified non-naturalistically, then self-referential incoherence threatens; if they are justified naturalistically, they will be unable to play any non-circular (or non-question-begging) role in the justification of naturalism."[16] He makes the same point against Roth's defense of Quine's naturalism, saying that: "This seems to me a very deep problem for the naturalized epistemologist: it must assume the legitimacy of, and strive to achieve, the very sort of justification it seeks to show cannot be had."[17] Naturalists themselves have considered the criticism seriously.[18]

I call the anti-naturalist counterargument 'a priori argument against methodological naturalism' – *a priori argument* for short. For assessing the argument, it is better to formulate it first. If we take (MN) for granted, then we can formulate the argument as follows:

> **Argument (AA)**: a priori argument against methodological naturalism
>
> 1. If (MN) is justified, it is justified either a priori or a posteriori.

[14] This expression has occasionally been used in slightly different meanings. See Audi, "Philosophical Naturalism," 42; Williamson, "What Is Naturalism?," 30.

[15] Audi, "Philosophical Naturalism," 42 (my italics).

[16] Siegel, "Naturalized Epistemology," 59.

[17] See Siegel, "Empirical Psychology," 675-676. See also Paul A. Roth, "Theories of Nature and the Nature of Theories," *Mind* 89, 355 (1980): 431-438; Roth, "Siegel on Naturalized," 485.

[18] See, e.g., David Papineau, *Philosophical Naturalism* (Oxford: Blackwell, 1993), 3-4.

2. A priori justification is not a proper justification.
3. If (MN) is (only) justified a posteriori, then its justification is viciously circular. (**VC**)
4. A viciously circular justification is not a proper justification.

∴ (MN) does not have a proper justification.

(AA) is a dialectical argument, since through it, it is shown that with espousing (MN), naturalists will be deprived of having a justification for their position and so, their position are totally unjustified.

All the premises of (AA) are supposed to be a priori. The premise (1) is based on the fact that every justification is exclusively either a priori or a posteriori. The premise (2) is a logical consequence of (MN) together with the assumption that a justification is proper only if it makes use of a method which can lead us to knowledge. The premise (4) seems to be endorsed by most epistemologists. These premises do not seem challenging. The main challenge of the argument is about the premise (3) – which I named (VC) – and the responses to the argument have mostly been accompanied with rejecting this premise. In the next section, I examine two responses to this argument and to the similar counterargument against the inductive argument for induction.

3. Two Responses to the Charge of Circularity

As I said before, some of the naturalists have considered the a priori argument seriously and responded to it. In this section, I examine two responses to the charge of circularity, one to the a priori argument (AA), and another to the similar objection to the inductive argument for induction.

3.1 Papineau's Response

David Papineau responds to the a priori argument in this way: traditionalists in that argument do not simply rest on a concept of philosophical activity; in fact, they have some serious presuppositions, the most important one is the assumption that knowledge should be certain, that is, it should be acquired via methods which *necessarily* deliver truth.[19] Applying the response to (AA), we can see that Papineau's main target is the premise (VC). According to (VC), if we employ an a posteriori method to justify (MN), we will be indulged in a vicious circle, since a posteriori methods are not legitimate in themselves, and they can only be employed with taking an assumption like (MN) for granted. So, traditionalists have the supposition that a posteriori methods are illegitimate until they are justified via a

[19] Papineau, *Philosophical Naturalism*, 4.

priori methods. But, Papineau rejects this supposition as unfounded. According to him, "[k]nowledge […] is the state that we need to get into if we are to succeed in avoiding error."[20] So, a belief-forming process (here, a method) to be "an effective means of avoiding error," should only be reliable in actual world, and not necessarily certain – that is, it needn't be "*impossible* for a given belief-forming process to produce a false belief."[21] Then, under this definition of knowledge, scientific methods should be regarded as legitimate in so far as they are reliable.[22]

Put it differently, there is a 'dialectical situation' here: each side of the dispute (naturalism vs. traditionalism) has its own methodology and the proponents of each side think that their own position should be justified via legitimate-according-to-them methods. Now, the important question is that "[…] which philosophical methodology should be used to address *this* issue [itself]?"[23] For Papineau, the response is to follow empirical methodology and the main reason behind his response is that "the onus surely lies with those who want to exclude relevant and well-confirmed empirical claims from philosophical debate to provide some prior rationale for doing so."[24] In fact, we can look at Papineau's opinion as though if two sides of a dispute have different viewpoints about the legitimate methods of inquiry, then none of them is allowed to criticize the other's position using her own method. Rather, if one's method *seems initially good enough*, the burden of proof is on the other side to show that her opponent's method is not legitimate (probably with the aid of a more reliable method).

I think that Papineau clarified the issue in the best way and I am in sympathy with his response to the a priori argument. But, I think, his response should be accompanied by a *meta-philosophical* explanation of the 'dialectical situation' and a suggestion to settle the dispute or, at least, a suggestion for how it could be judged. Therefore, although Papineau's response has many things in common with the proposal I am going to put forward in this paper, my proposal, I think, will be more acceptable since it also contains a meta-philosophical explanation of the situation and a suggestion for settling the dispute.

---

[20] Papineau, *Philosophical Naturalism*, 143.

[21] Papineau, *Philosophical Naturalism,* 144 (original italics).

[22] For more details, see Papineau, *Philosophical Naturalism*, 143-152.

[23] Papineau, *Philosophical Naturalism*, 4 (original italics).

[24] Papineau, *Philosophical Naturalism*, 4. See also *Philosophical Naturalism*, 156-157; Rosenberg, "Can Naturalism Save," 39.

3.2 Baithwaite-van Cleve Response

In addition to the a posteriori argument for methodological naturalism, some of the other *naturalistic arguments* which have been put forward to justify the reliability of specific naturalistic methods have been accused to the charge of circularity as well. Richard Fumerton, for instance, accuses an inductive argument for justifying induction (and any other argument of this sort) of circularity. Based on his analysis, he rejects epistemological externalism.[25] Another example is Richard Boyd's naturalistic argument in defense of scientific realism which is known as *explanationist* (or abductive) *defense of realism* (EDR).[26] The argument makes use of inference to the best explanation (IBE) to show that IBE is a reliable method: the best explanation of the success of IBE in science is that it is reliable. Then, using IBE itself, IBE is reliable, and therefore, what our best scientific theories (which are justified using IBE) tell us is nearby true. This argument has been accused of circularity by Arthur Fine on the ground that it makes use of the very method which the argument is supposed to justify.[27]

Richard Braithwaite and James van Cleve separately tried to respond to the potential-in-their-times (later, an actual) accusation of circularity against a naturalistic inductive argument for the reliability of induction.[28] Their response is simply as follows: an argument can be circular in two different ways: 'premise-circular' and 'rule-circular.' An argument is premise-circular if its result is contained in (at least) one of its premises. This kind of circularity is apparently vicious. But, if

---

[25] Richard A. Fumerton, "Skepticism and Naturalistic Epistemology," *Midwest Studies in Philosophy* 19, 1 (1994): 321-340, 337-338; Richard A. Fumerton, *Metaepistemology and Skepticism* (Maryland: Rowman & Littlefield Publishers, 1995), 177,180. See also Laurence BonJour, *Epistemology: Classic Problems and Contemporary Responses*, 2nd ed. (Maryland: Rowman & Littlefield Publishers, 2010), 64.

[26] See Richard N. Boyd, "On the Current Status of the Issue of Scientific Realism," *Erkenntnis* 19 (1983): 45-90; Arthur Fine, "Unnatural Attitudes: Realist and Instrumentalist Attachments to Science," *Mind* 95, 378 (1986): 149-179; Stathis Psillos, *Scientific Realism: How Science Tracks Truth* (London: Routledge, 1999), chap. 4.

[27] See Arthur Fine, "The Natural Ontological Attitude," in *Scientific Realism*, ed. Jarrett Leplin (Berkeley, CA: University of California Press, 1984), 83-107, esp. 85-86; Fine, "Unnatural Attitudes."

[28] See Richard Bevan Braithwaite, *Scientific Explanation: A Study of the Function of Theory, Probability and Law in Science* (Cambridge: Cambridge University Press, 1953); James van Cleve, "Reliability, Justification, and the Problem of Induction," *Midwest Studies in Philosophy* 9, 1 (1984): 555-567. Psillos (*Scientific Realism*) follows their approach to respond to the accusation of circularity against EDR.

an argument proceeds with the same method which the argument is aimed to justify, it is rule-circular and such circularity is not vicious at all.[29]

Now, if we are to use this strategy to respond to the charge of circularity against the a posteriori argument for methodological naturalism, we should say that the argument makes use of empirical methods (the rules) to justify the legitimacy (and superiority) of those methods in philosophical research. Therefore, it is rule-circular rather than premise-circular, and viciousness of rule-circularity is far from being intuitive. Hence, the premise (VC) of the argument (AA) is false. So, the a posteriori argument for methodological naturalism is good enough to be accepted as a justification for methodological naturalism.

But, this response does not seem desirable and it has been criticized even by some naturalists. As an example, James Ladyman and Don Ross criticize preserving EDR with the aid of a similar tactic saying that even if that argument is not premise-circular and then, not viciously circular, "this style of argument will not persuade someone who totally rejects IBE." This argument can only show that using IBE in philosophy of science is 'consistent' and can be part of an 'adequate philosophy of science.'[30] Igor Douven also criticizes this tactic saying that if we allow rule-circularity, some improper reasoning rules which can be forged so that to be able to be justified using themselves will be permitted as well. For example, 'Inference to the Worst Explanation' (IWE) as an imaginary rule may be defended on the grounds that it leads to the most unsuccessful theories from the available data. Then, using IWE itself, the worst explanation of the failure of those theories is that they are true. However, it is completely irrational to conclude that IWE is a reliable rule of inference leading us to correct theories. Furthermore, Douven thinks a rule-circular argument for IBE has just the power to persuade the ones who have inclination to use IBE, not the ones who totally reject it – as Ladyman and Ross mentioned too.[31]

---

[29] Braithwaite, *Scientific Explanation*, chap. VIII; van Cleve, "Reliability, Justification," 557-559; Papineau, *Philosophical Naturalism*, 154-158.

[30] James Ladyman and Don Ross, "Scientific Realism, Constructive Empiricism, and Structuralism," in *Every Thing Must Go: Metaphysics Naturalized*, ed. James Ladyman, et al. (Oxford: Oxford University Press, 2007), 66-129, 75.

[31] See Igor Douven, "Abduction," *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta (2011), http://plato.stanford.edu/archives/spr2011/entries/abduction/ (accessed 18 August 2016), sec. 3.2.

4. Reflexivity: A Requirement for Self-Referential Theories

As I tried to show in the previous section, the naturalists' responses to the a priori argument against methodological naturalism do not seem so compelling, especially when considering the situation meta-philosophically. In this section, I put forward a proposal which, I believe, shows that the a priori argument against methodological naturalism is totally unfounded. For this reason, I focus on the requirements a philosophical theory should meet to be a candidate for solving a philosophical issue or problem in some area of philosophy or other. Generally speaking, it seems satisfactory to say that every philosophical theory should fulfill some requirements as prerequisites of being a genuine philosophical theory. A similar idea has been posed and defended about scientific theories by Larry Laudan who called such conditions 'methodological rules or standards' and tried to justify them.[32] To make it analogous with this caption, I call such a requirement for philosophical theories 'meta-methodological rule or requirement for philosophical theories' (hereafter MMR). The most general and obvious MMR – which has a parallel in science – is the requirement that every philosophical theory should be *internally consistent*. Even if some acceptable-at-face-value philosophical theories actually do not meet internal consistency requirement, there is a consensus, I think, that after their inconsistency are revealed such theories should be made consistent to be proper philosophical theories. This MMR does not seem challenging. About scientific theories, Laudan even mentions some quotations from relativists like Thomas Kuhn who endorse the consistency requirement for scientific theories.[33]

Now, I want to introduce another MMR – though it is not as general and pervasive as consistency requirement – which can be called *reflexivity requirement*. Before formulating that MMR, I should introduce three other concepts. First, *target domain* of a philosophical theory is the set of entities or n-tuple of entities which that theory is aimed to explain, reduce or unravel something about them or the relation between them – in the case of n-tuples with $n>2$.[34] As a scientific example, Darwinian theory of evolution with natural selection is about different *organisms* to explain such things as their variation, their evolutionary history and so on.[35] Therefore, its target domain is the set of all organisms in all times and places. In philosophy of mind, type identity theory is a familiar example. If we understand the

[32] Larry Laudan, *Beyond Positivism and Relativism: Theory, Method, and Evidence* (Colorado: Westview, 1996), chap. 5 and *passim*.

[33] Laudan, *Beyond Positivism*, 92.

[34] Target domain of a theory can be said to be the same thing as what Fitch calls 'subject-matter' of a theory (Frederic B. Fitch, "Self-Reference in Philosophy," *Mind* 55, 217 (1946): 64-73, 64).

[35] Fitch, "Self-Reference," 64.

theory as stating 'every mental property is identical with some physical property,' then the target domain of this theory is 2-tuples of different mental and physical properties as (M_i, P_i), and the theory is about the relation between the 2-tuples entries. And if we consider the theory as stating '*necessarily*, every mental property is identical with its related physical property,' then we should consider the 2-tuples across all possible worlds, and the target domain of the theory becomes a set like:

$\{(M_1 \text{ in } w_1, P_1 \text{ in } w_1), (M_2 \text{ in } w_1, P_2 \text{ in } w_1), \ldots, (M_1 \text{ in } w_2, P_1 \text{ in } w_2), \ldots\}$.

Second, *actual domain* of a theory is a subset of its target domain which the theory is true about its members. In practice, every theory is proposed by its theoretician as a true theory about all members of its target domain. But, it is just an ideal for every theory and not necessarily the actual case. In fact, thought experiments in philosophy are mostly aimed (as some tools) to show that a theory is not true about some (or even all) members of its target domain, or to put it another way, to show that its actual domain is not identical to its target domain. However, if actual domain of a theory is identical to its target domain, then that theory, at least at first sight, is successful.

Third, I define a *self-referential theory* as a theory that is itself directly or indirectly within its target domain (Fitch's similar definition will be introduced in the next section). I will say more about this sort of theory and its significance in philosophy in the following.

Now, I define reflexivity as a property of some philosophical theories: a theory, T, is reflexive, if and only if

> (**RC**): if T (directly or indirectly) is a member of its target domain (i.e. if it is self-referential), then it is a member of its actual domain as well.[36]

To explain (RC), I start with a trivial example. Suppose, for the sake of argument, that the claim 'all theories are false' is a philosophical theory (call it (F)). The target domain of (F) contains all theories in all areas of science, mathematics, and philosophy. Suppose (in sympathy with scientific realism) there are some true scientific theories, or instead, suppose there are at least some true mathematical theories. Then, (F) is clearly false. It is because the actual domain of (F) is not identical with its target domain. But, (F) – regarded as a theory – is itself *directly* a member of its own target domain and hence self-referential, and since it is false, it is a member of its actual domain as well. So, it is reflexive. If we change (F) into the claim 'all propositions are false' (call it (F')), then *the proposition stating* (F') will be

[36] Cf. Fitch, "Self-Reference."

a member of the target domain of (F'). So, (F') is *indirectly* self-referential.[37] According to many true propositions – from the trivial ones like 'every table is a table' to even some scientific laws – (F') and its relevant proposition are false. Therefore, it is reflexive.

Now it seems to me that reflexivity requirement, *the condition that every self-referential theory should be reflexive and fulfill (RC)*, is a genuine (though special) MMR.[38] The main reason behind the claim is that we expect every theory to embrace all the members of its target domain, and so, if that theory is a member of its own target domain (directly or indirectly), it should be true about itself as well as other members of its target domain.[39]

Reflexivity requirement so described does not seem challenging and, I think, most philosophers will endorse it. It also is not a trivial requirement. To show this, I put forward two cases in philosophy that can be regarded as evidence of upholding reflexivity requirement by some philosophers (implicitly or explicitly) as a necessary and nontrivial condition for philosophical self-referential theories.

4.1 Logical Positivists' Verification Principle: Charge of Self-Refutation

The verification principle of logical positivists was asserted by them as a criterion of meaning for sentences of any language[40], and it was posed in different forms, each of which was an endeavor to overcome the difficulties of the previous versions. In its simplest form, the principle states that a sentence is cognitively meaningful if it is analytic (or self-contradictory) – i.e. its truth or falsity can solely be judged on the meaning of its terms – or it is (at least in principle) verifiable – that is, its truth or falsity can be verified with empirical tests.[41] The principle has been criticized in different ways, one of which is intended to be discussed here: the charge of being

[37] I do not assume that a theory is combined of one or more propositions. Rather I assume, for every theory, there is a proposition (or a set of propositions) which *states* that theory. In comparison, Fitch only mentions the direct way of "self-reference." Therefore, to regard the theories concerning propositions as self-referential, he supposes, following Whitehead, that all theories are propositions and all propositions can be seen as theories (see Fitch, "Self-Reference," 65).

[38] According to (RC), if a theory is not contained in its target domain – i.e. if it is not self-referential – it is trivially reflexive, because the antecedent of (RC) is false, and so, (RC) will be true about the theory.

[39] Cf. Fitch, "Self-Reference."

[40] The criterion was initially posed for sentences. But, after a while, positivists focused on propositions rather than sentences. Here, I considered the initial version of the criterion.

[41] See, e.g., Carl G. Hempel, "The Empiricist Criterion of Meaning," in *Logical Positivism*, ed. Alfred Jules Ayer (New York: The Free Press, 1959), 108-127, 108.

self-refuting or *self-undermining*. The criticism is leveled at the principle through a question: if we apply the criterion to itself, does it itself meet the criterion, i.e. is it cognitively meaningful under itself? If it is, it should be either analytic or verifiable. If it is analytic, then it is an arbitrary principle and an empty judgment. And it cannot be verifiable regarding any concept of verifiability.[42]

Although it is stated that verificationists did not consider the challenge seriously,[43] some of the main figures of the position, including Rudolf Carnap, A. J. Ayer and Carl Hempel, met the challenge and tried to respond to it.[44] I think the criticism is relevant and the best way to explain it and the tries carried out to respond to it is via reflexivity requirement: verification principle as (part of) a theory of meaning is indirectly (via the sentence which states the principle) self-referential, and therefore, should be able to apply to itself. This fact *explains* the self-undermining charge to verification principle and the verificationists' tries to respond to it.

4.2 Strong Programme in Sociology of Scientific Knowledge

One of the main currents in Sociology of Scientific Knowledge (SSK) is the so-called Strong Programme. According to the school of thought, science should be studied by sociologists of science as a 'natural phenomenon' and as a culture or collective belief of some people (scientists, students etc.) who work within a community (the scientific community). In such studies, *whole* science as the subject matter of sociologists of science should be scrutinized rather than the only parts of science which are mostly considered false and unsuccessful; just like physiology which its aim "is to explain the organism in [both] *health and disease*."[45] Following this approach, sociologists of science do study the social, economic, and cultural conditions in which a scientific theory emerged and was accepted at a specific time in the (close or distant) history of science, and their effect on the emergence and acceptance of that theory. In brief, for proponents of Strong Programme, SSK should

---

[42] See, e.g., Hilary Putnam, *Reason, Truth and History* (Cambridge: Cambridge University Press, 1981), 106 and *passim*. Putnam tries to show that all the viewpoints like verificationism are subject to the same criticism. I think Putnam's analysis can be reconstructed with the aid of reflexivity requirement I described here.

[43] Putnam, *Reason, Truth,* 106.

[44] See Rudolf Carnap, *Philosophy and Logical Syntax* (Bristol: Thoemmes Press, 1996), 36-38; Alfred Jules Ayer, *Language, Truth, and Logic* (New York: Dover, 1952), 16; Alfred Jules Ayer, ed. *Logical Positivism* (New York: The Free Press, 1959), 15-17; Hempel, "The Empiricist Criterion of Meaning," 123-126.

[45] David Bloor, *Knowledge and Social Imagery*, 2nd ed. (Chicago: University of Chicago Press, 1991), 5 (my italics).

aim to discover the social 'causes' of scientific theories and their acceptance. Social factors are not limited to some forces from outside the scientific communities. Rather, there are many social factors within a scientific community.[46]

According to the above characterization of Strong Programme in SSK, the idea itself can be seen as a theory or approach in sociology aimed to study a 'natural phenomenon,' i.e. science. So, it is subject to this potential criticism that if we endorse it as the *correct* theory about science, then its own emergence and the expansion of its proponents must also have some social causes and social explanation. David Bloor paid attention to this criticism. But, before responding to it, Bloor had gone further introducing 'reflexivity' as one of the tenets of Strong Programme. He says about Strong Programme that: "In principle its patterns of explanation would have to be applicable to sociology itself," and he sees this requirement obligatory "because otherwise sociology would be a standing refutation of its own theories."[47] We can paraphrase Bloor's words with saying that neglecting reflexivity requirement leads Strong Programme (which is self-referential) to be excluded from its actual domain.

Bloor also takes this possibility into consideration that reflexivity requirement can be a threat to his viewpoint as an argument showing that the viewpoint is *self-refuting*. Such an argument proceeds in this way: if Strong Programme is reflexive, then part of its cause is related to social factors, and it makes Strong Programme unjustified and false. Bloor narrates some forms of this criticism from Grünwald's and Lovejoy's works.[48] Regardless of the details of Bloor's response to the charge of self-refutation, his response does not include any denial of reflexivity of Strong Programme.

5. Reflexivity vs. Consistency

There is a potential serious objection to the reflexivity requirement here.[49] The objection is that, after all, reflexivity requirement is just a special case of consistency requirement, and then, it should not be posed as an independent requirement of philosophical theories. But, I do not think it is right. I give two complementary reasons for why I think reflexivity is (and should be regarded as) an independent requirement, though the objection will not threaten my response to the a priori argument against methodological naturalism.

---

[46] See, e.g., Bloor, *Knowledge and Social*, 5-7.

[47] Bloor, *Knowledge and Social*, 7.

[48] Bloor, *Knowledge and Social*, 17-18,44.

[49] Kim Sterelny and Hossein Sheykh-Rezaee reminded me about this objection.

The first reason is that reflexivity is not the same property of theories as consistency, because a theory may be both reflexive and *inconsistent* at the same time. As an instance, the quasi-theory (F), the theory that all propositions are false, is inconsistent but reflexive.

However, it is not a plausible response to the criticism, because the reverse case is not true: a theory cannot be *irreflexive* but consistent. In fact, the main criticism is that we do not need reflexivity as an independent requirement, and consistency requirement is enough to be satisfied by theories. So, I need to give the other reason.

The second reason is not theoretical but a practical one. It seems that reflexivity is usually regarded as an independent requirement, or at least as an important special sort of consistency requirement worth mentioning as a separate requirement. If a theory is not reflexive, it is not usually said that it is inconsistent. Rather, such a theory is usually said to be *self-refuting* – as was the case about logical positivists' Verification Principle and Strong Programme. Although a self-refuting theory is inconsistent, it is inconsistent *in a special sense*; i.e. holding such a theory to be true entails its negation. A theory which is not within its target domain is not subject to self-refutation. A theory which does not have any other theory within its target domain is, a fortiori, not subject to self-refutation as well. Roughly speaking, Frederic Fitch calls the latter a 'horizontal' theory and other sort of theories 'vertical.'[50] Contrary to scientific inquiry, vertical theories are not rare in philosophy. As Fitch correctly says, in philosophical research,

> extreme comprehensiveness is sought for. Theories are constructed which purport to deal with all entities whatsoever and which therefore have an unrestrictedly extensive subject-matter. In dealing with all entities, such theories in particular deal with all theories, since theories are themselves entities of a special sort. In philosophy we thus encounter theories about the general nature of theories.[51]

He calls a vertical theory "included in its own subject-matter" – that is, in its own target domain – 'self-referential.'[52] According to him, only self-referential theories may be self-refuting, or in his words, 'self-referentially inconsistent.' He says that a horizontal theory

> may be internally inconsistent, or it may be inconsistent with known facts, and hence 'externally' inconsistent, but it cannot be inconsistent with its own nature

[50] In fact, his definitions of vertical and horizontal theories are more complicated. He introduces the concept "ordinal level" of a theory to define vertical and horizontal theories. For the exact definitions, see Fitch, "Self-Reference," 64-65.

[51] Fitch, "Self-Reference," 64-65.

[52] Fitch, "Self-Reference," 65.

> in the way that a self-referential theory can. If a self-referential theory T implies that T has the property P, and if T in fact does not have the property P, then we shall call T self-referentially inconsistent.[53]

Self-referentially inconsistent theories are, in fact, those self-referential theories which do not satisfy reflexivity requirement. So, as Fitch shows, it seems satisfactory to consider reflexivity to be a special sort of consistency. But, why should it be regarded as an independent requirement apart from consistency? The answer to this question, I think, lies in the *function* those requirements fulfill in theory appraisal.

First, to find out what the function of consistency requirement in theory appraisal is, we should understand why theories should satisfy that requirement. It seems that consistency requirement provides us with a minimal rationality: it is not rational to endorse a theory which may imply two contrary consequences. Now, I think that reflexivity requirement has a different function in theory appraisal that makes it worthy of being an independent requirement. As said above, reflexivity requirement is mainly for self-referential theories; theories like relativism,[54] positivism, some kinds of skepticism, social constructivism, any analysis of language (which essentially makes use of language itself), any theory of truth and meaning (which itself essentially bears truth value and meaning[55]), conspiracy theory, and more general and more important, any meta-philosophical doctrine about philosophy and philosophical inquiry. Such theories have usually been influential and permeating and they constitute an important part of philosophy and probably most of meta-philosophical literature. Some of these theories – like relativism and positivism – have encountered accusation of irreflexivity in practice. Sometimes, the response has been to evade reflexivity requirement with excluding the theory in question from its target domain; that, as far as I am concerned, is an *ad hoc maneuver*. Consistency requirement is too general to prevent theorists from carrying out this ad hoc maneuver in practice. There should be a more specific requirement to attract theorists' attention to avoid such a maneuver when developing their theories. Reflexivity requirement can fulfill this function. Following the requirement, a theorist will be aware that she should develop her self-referential theory so that it includes itself in its actual domain, or her theory should not be self-referential at all, albeit she should not appeal to an ad hoc maneuver to exclude her theory from its target domain.

---

[53] Fitch, "Self-Reference," 66. The property P may be said to be the property that T ascribes to all the theories in its target domain.

[54] See Maria Baghramian, *Relativism* (London: Routledge, 2004), 100-107.

[55] See, e.g., Roth, "Theories of Nature."

6. Methodological Naturalism and Reflexivity Requirement

Now, if we see methodological naturalism as a philosophical theory expressed through (MN), what is its target domain? To answer the question, it should be made clear what (MN) is about. (MN) prescribes following scientific method for philosophers. But, what should we do if we are to assess an *existing* piece of work or a theory in philosophy to see whether and to what extent it is naturalistic in this sense? To put it another way, what works in philosophy do naturalists endorse as naturalistic? If naturalists are asked to name some naturalistic works in philosophy, they will definitely cite, inter alia, evolutionary works (in different areas of philosophy like epistemology, ethics and philosophy of science) as their paradigm cases of naturalistic works in philosophy.[56] Now, what does it make such theories naturalistic? It does not seem enough for a thoroughgoing naturalistic work to just be inspired or informed by a scientific theory. For if, say, evolutionary epistemologists justified their view *traditionally* using armchair methods, their view would be no longer considered to be naturalistic. What makes, say, a theory in evolutionary epistemology of theories (EET) naturalistic is, I think, the assertion that it is (as claimed) *confirmed* with relevant evidence of the *actual* changes of theories in the past history of science and it will probably be confirmed in the future as well – though under a general sense of 'confirmation.' It means that EET to be naturalistic should be justified abductively, using IBE as a scientific method.

If I am right, it seems that (MN) can be rewritten as follows:

> (**MN′**): Any philosophical theory can constitute knowledge only if it is justified scientifically (i.e. through scientific method).[57]

According to (MN′), the target domain of methodological naturalism is the set of all philosophical theories within which, there is methodological naturalism itself, and so, it is self-referential.[58] Hence, methodological naturalism to be reflexive should be true about itself, that is, it should be justified through scientific method. Therefore, not only does an a posteriori argument for methodological naturalism not lead to a vicious circle, but also it shows that methodological naturalism fulfills reflexivity requirement through such an argument.

There are some clues in the works of anti-naturalists about this matter. In fact, it seems that some anti-naturalists take reflexivity requirement (implicitly) for

---

[56] See Rosenberg, "A Field Guide."

[57] See Elliott Sober, *Ockham's Razors* (Cambridge: Cambridge University Press, 2015), 245.

[58] Fitch makes a similar point when he says: "consider the view that every valid theory must be obtained from observed empirical data. This is a theory about theories and their validity" (Fitch, "Self-Reference," 66).

granted and try to show that methodological naturalism is not reflexive since it can only have an a priori justification if it is justified at all. The criticism can be traced back to Sheldon's 1945 article. The quotation below is one of his attacks to '1944 school of naturalism:'[59]

> Your own creed tells you not to believe anything till it is experimentally confirmed. How can you expect us to believe you have the right method for philosophy until you show us that it succeeds in giving us objective truth comparable with that of the sciences, truth on which the philosophic experts agree? […] I fear that you claim to have proved the rightness of your method in philosophy beforehand; an a priori claim you should be the last to make.[60]

As is immediately apparent in the last sentence of the above quotation, Sheldon states that naturalists to be coherent (better to say, reflexive) are not allowed to appeal to a priori defense of naturalism. This is the first horn of naturalist's dilemma.

Williamson also drops a similar hint saying that this claim of radical naturalists that "all truths are discoverable by hard science" is expected to be discoverable through hard science. But it is not, Williamson claims, and so, this kind of naturalism is not true.[61]

Fitch also touches on the point. About "the view that every valid theory must be obtained from observed empirical data," he writes:

> Incidentally it is a theory which does not seem to conform to its own criterion as to what constitutes a valid theory, at least not unless it can itself be shown to have been obtained as a generalization from observed empirical data.[62]

At the end of this section, another point should be added. The methodological traditionalism is also reflexive because it is justified through traditional methods. So, the main point remains here is that how we can compare these two methodologies in philosophy; the issue that will be addressed in the next section.

7. Reflexivity and Justification

What I have tried to show up to this point is that having an a posteriori argument for methodological naturalism does not make it viciously circular at all, and rather, according to reflexivity requirement for self-referential theories, such an argument for naturalism is mandatory and satisfying the requirement should be considered a

[59] See Kim, "American Origins," 86; Sheldon, "Critique," 253.
[60] Sheldon, "Critique," 268.
[61] Williamson, "Unclarity," 37. See also BonJour, "Against," 7.
[62] Fitch, "Self-Reference," 66. See also Siegel, "Naturalized Epistemology," 57-58; Siegel, "Empirical Psychology," 675-676.

merit for methodological naturalism, and then, traditionalist a priori argument against methodological naturalism is totally unfounded.

But now, there remains a question here: does reflexivity of a theory confer justification on it? I do not suppose it does. The main reason behind this is the fact that being reflexive means that a theory is true about one (though a special) member of its target domain, not all (or even most) of the members of its target domain comprehensively. Therefore, being reflexive does not confer an *adequate* justification on any theory and a theory should be justified independently of being reflexive.

But the problem with methodological naturalism is specifically thornier. It is because of the fact that any kind of justification proceeds following a special method. And methodological naturalism is a meta-philosophical idea (or norm) about the legitimate methods of philosophical inquiry. Now, which method should be adopted to justify a methodological view like methodological naturalism? Papineau, tackling with the charge of circularity, drops a hint about this matter: such an accusation can be leveled against traditionalism since it is a methodological view as well.[63] In fact, we have two competing views each of which has its own methodology, and as a result, each prescribes a kind of justification according to its own method. So, the dispute apparently ends in stalemate and we are left with no choice but to admit those views are *incommensurable in a sense*, i.e. methodologically incommensurable.

Incommensurability of scientific theories has been a problem after the works of Thomas Kunh and Paul Feyerabend who infamously known as the founders of modern relativism. But there are some solutions to this problem, posed by philosophers of science like Imre Lakatos and Laudan. Now, if the problem about scientific theories has some (for me, compelling) solution, why not try it on the parallel problem about philosophical theories? I believe that the best solution to the problem of incommensurability about scientific theories (which I think is Laudan's) can in the same way be employed to address the problem of incommensurability of philosophical methodologies. There is not enough space here to elaborate the view, but I introduce the main elements of the suggestion.

Laudan in addressing issues like scientific progress and incommensurability of scientific theories in history of science, proposed his concept of *research tradition* – similar to but not identical with Kuhn's concept of 'paradigm.' A research tradition is a more general view than a single theory, that includes ontological and methodological parts and overarches theories which follow the ontology and methodology of the tradition. Therefore, two theories from two different research

[63] See Papineau, *Philosophical Naturalism*, 157-158.

traditions even differ on their methods. Such theories cannot (and should not) be compared with each other directly. Instead, we can compare their traditions (even if they are incommensurable considering their basic ontological entities and/or their research methods). Competing traditions are compared according to their *problem-solving effectiveness*, that is, the total number of problems solved by (the theories of) each of them (with weighing their importance).[64] Problem-solving effectiveness of a tradition shows the success (or failure) of it: the tradition with more amount of problem-solving effectiveness – that which solves more problems (and the most important ones) – is more successful. Progress rate of a tradition can be measured by the problem-solving rate of that tradition, and we should also invest in traditions which show high progress rates awhile after their appearance.[65]

This *model* works in philosophy as well. Such views like naturalism and traditionalism, I think, are not philosophical theories or theses; rather, they are research traditions. Naturalism, for example, has its own methodology and ontological constraints (the latter due to *ontological* naturalism), and many philosophical theories follow its methodology and ontology. Therefore, naturalism and traditionalism should be assessed and compared to each other according to their success in problem-solving. If this model is pertinent, the only way for justifying naturalism (and traditionalism as well) is appraising its success in problem-solving in comparison with its rival, rather than trying to adduce a philosophical argument for (or against) it. I hope to articulate this view in another occasion.[66]

---

[64] In measuring the problem-solving effectiveness of a research tradition, we should also take into account the *conceptual problems* the tradition poses. The number of conceptual problems (considering their importance) decreases the problem-solving effectiveness of a tradition (Larry Laudan, *Progress and Its Problems: Towards a Theory of Scientific Growth* (Berkeley, CA: University of California Press, 1977), chaps. 2-3).

[65] For more details, see Laudan, *Progress*. For a brief introduction to his proposal, see Laudan, *Beyond Positivism*, chap. 4.

[66] The research project that led to this article was funded by Iranian Institute of Philosophy. Part of the research has been carried out at the Australian National University. I should thank the officials of ANU School of Philosophy and Research School of Social Sciences for letting me stay at the quiet office which enabled me to conduct the research. I am deeply indebted to Kim Sterelny, Hossein Sheykh-Rezaee, and Laleh Ghadakpour, who lend me a great deal of support in my research project. I should also thank Alexander Rosenberg, Alexander Sandgren, and Stephen Mann for their valuable comments on earlier drafts of this paper.