

Forthcoming in **Philosophical Studies**

---

# Transcendental Idealism among the Jersey Metaphysicians

Gordon Belot

Received: date / Accepted: date

I too lack sympathy for metaphysics,  
though not in general: only for  
pre-Kantian metaphysics—and then  
only if practised after Kant.

---

van Fraassen

## 1 Introduction

Let me begin with a conceit—unworthy of Bas, granted, but intended as a tribute nonetheless.

\*\*\*

*An excerpt from a grant proposal of the distant future.* There has recently been a revival of interest in the so-called Jersey school of metaphysicians. The members of this school flourished long ago of course, in the tumultuous period immediately preceding the biochemical amalgamation of the Bush and Clinton dynasties. Yet their views and arguments remain of contemporary interest.

Recent literature on the Jersey metaphysicians has provided a portrait of the movement in broad brushstrokes, emphasising the substantial body of doctrinal agreement within the school, with a special focus on their peculiarly rabid form of transcendental idealism. I aim to fill in some of the details in this portrait, by exploring some of the differences of detail between the views of the two influential early Jersey metaphysicians, David Lewis and Bas van Fraassen.<sup>1</sup>

---

Gordon Belot  
Department of Philosophy  
University of Michigan  
E-mail: [belot@umich.edu](mailto:belot@umich.edu)

<sup>1</sup> Although all undergraduates today associate these philosophers with the doctrines known as *Lewis's Beard* and *van Fraassen's beard* (the one so much easier to grasp than the other!), there is little reason to think that either actually espoused the doctrine now associated associated with his name.

It is clear from the doxography that during the period in question Kant was the dominant figure of study in Jersey.<sup>2</sup> What set Lewis and van Fraassen apart from some of their contemporaries was their shared focus on the metaphysics of transcendental idealism. As one can learn from any textbook today, each of these authors built his system around the thesis that the fundamental structure of the world, as discoverable through experience and science, is not objective but has a certain relativity to cognitive agents.

I have explored Lewis's Kantianism in previous work. At least two of Lewis's central doctrines are Kantian in provenance. According to the first, due to the nature of the principle of individuation of properties the reach of science is limited and cannot provide us with knowledge of the intrinsic nature of things-in-themselves (Lewis 2009, Langton 2004). According to the second—which provides a clear example of transcendental idealism in our sense—the causal structure of the world is not absolute but rather depends on our cognitive constitution: the causal relations of our world are determined by which counterfactual statements are true (Lewis 1986a); which counterfactual statements are true depends on the laws of nature (Lewis 1986b); and the laws of nature depend on our cognitive capacities (Lewis 1986c, Appendix C).<sup>3</sup>

In my future work I intend to explore the transcendental idealism of van Fraassen, by focussing on two of his works: an early work on the philosophy of space and time (van Fraassen 1970); and a work of his middle period concerned with scientific representation (van Fraassen 2008).<sup>4</sup> My preliminary studies suggest the following picture. In the early work, van Fraassen espouses a variety of transcendental idealism according to which the structures of space and time are relative to our scientific beliefs, so that there is a sense in which reality, in abstraction from the beliefs of agents, has no spatial or temporal structure—in particular, the question of “correctness” of a theory about the structure of space and time does not arise. The discussion of transcendental idealism in the later work is, curiously, not nearly so explicit. But many of the elements of the earlier work seem to still be in place. So it would appear that for van Fraassen of the middle period, the early work erred only in not stressing that the basic structure of the argument was extremely general, and did not depend on any special feature of the philosophical problems of space and time. If all of this is correct, then an interesting picture emerges, on which van Fraassen's transcendental idealism differed in important ways from that of Lewis—and was, if anything, an even more bold piece of metaphysics.

\*\*\*

---

<sup>2</sup> Note that recent scholarship has called into question the traditional attribution of the panegyric “Kant über alles” to Harman *père*.

<sup>3</sup> Permit me an observation that will illustrate the caprice of the processes of history by which writings from that period were transmitted to us. We know that van Fraassen advanced an account of the nature of science upon which this aim was relative to human perceptual capacities. We know that his contemporaries found this view curiously difficult to accept and often asked him to explain how this account was compatible with the fact that human perceptual capacities might change, through technology or through redrawing of the human/nonhuman boundary. Surely Lewis's more radical position was met with the same sort of Boetian uproar. Yet no trace of this remains in the works that have come down to us. Equally curious: of the two Lewisian theses just mentioned, only the first is characterised as Kantian in the extant literature of that time.

<sup>4</sup> I aim eventually to develop the sort of expertise in the history of trapeze and of long-vanished varieties of wine and cheese that would allow me to grapple with van Fraassen's notoriously intricate late works. Unfortunately this will not be feasible in the temporal framework of the present project.

Well, that is more than enough of that. But I hope my intentions are clear at any rate. In the preface to his sparkling recent book, *Scientific Representation: Paradoxes of Perspective* (van Fraassen 2008, henceforth ‘*SR*’), van Fraassen intimates that neo-Kantian themes will be evident throughout the book. And they are, especially in his discussion of the problem of coordination (Chapter 5) and of the prospects for empiricist structuralism (Chapter 11). But I want to use this occasion to press van Fraassen to make these themes more evident still.

In particular, I would like to understand better the relation between the views developed in the new book, and those developed in his book on the philosophy of space and time: the notion of a logical space appears to play a similar (and similarly prominent) role in both works—and in the older book this notion was intimately bound up with the striking anti-realism about spatial and temporal structure that van Fraassen then espoused. Further, I would like to understand whether new work can be read as a contribution to metaphysics, rather than a recusal from it.

In the final section of the paper I pose some questions for van Fraassen. The intervening sections provide context and motivation.

## 2 Good-Cop, Bad-Cop

In some moods, van Fraassen evinces a tolerant attitude towards the sort of ontological investigations that are at the centre of so much of contemporary metaphysics and philosophy of physics. In a recent discussion, he characterises the project of interpreting physical theories in the following terms.

The enterprise of interpretation, its flagships being twentieth century interpretations of relativity theory and quantum mechanics, is not a pursuit of truth. The nearest I can say is that it pursues a sense of understanding in which the question of truth is bracketed, in order to give us a handle on the conceptual structure of a theory, seen from various perspectives, and to give us some assurance of its coherence (van Fraassen 2007, p. 358).<sup>5</sup>

Jauernig suggests to van Fraassen that we should think of metaphysics as having a corresponding role:

to provide theories that complement our scientific theories about the world and the human condition in order to arrive at a more complete story . . . . These metaphysical theories are not presented as true but as metaphysically adequate, that is, as possible and explanatorily or intellectually satisfying. . . . [I]sn’t it plausible to say that understanding the world and the human condition involves knowing the *possible* ways the world could be with respect to those aspects or domains that cannot in principle be empirically investigated, and which are, thus, inaccessible to the empirical sciences? (Jauernig 2007, pp. 315 f.)

Responding to Jauernig, van Fraassen suggests that his own preferred account of metaphysics differs from hers only in tone (van Fraassen 2007, §5). Presumably, for this tolerant van Fraassen, the only general complaint to be made against metaphysics as currently practised is an accusation of false consciousness—too many metaphysicians believe falsely that the aim of their field is to produce true theories. But there is a

<sup>5</sup> See also (van Fraassen 1989, §9.4) and (van Fraassen 1991, §§1.1, 1.3, 8.1, and 12.5).

sense in which this is a minor criticism: for on van Fraassen's view many scientists labour under an analogous misapprehension—but that fact does little to detract from the glory of science.

But van Fraassen also has his sterner moods. In another recent discussion, he investigates the prospects for a *rapprochement* between empiricism and neo-Kantianism (van Fraassen 2009). He identifies as an obstacle to *détente* the fact that each school sees the other as mired in metaphysics (p. 459)—the empiricist sees the neo-Kantian as embracing a dependence of reality on the agent, the neo-Kantian sees the empiricist as naively embracing a species of realism—while itself taking metaphysicians to err not in espousing false theses, but in failing to espouse any at all (p. 474). In the course of this discussion, van Fraassen labours mightily to show that the account of scientific representation that he offers in Parts II and III of *SR* should not be understood as just another item on the metaphysicians' menu, but rather as opting out of the presuppositions that shape the contemporary debate concerning the metaphysical underpinnings of intentionality (van Fraassen 2009, §§3.4 and 4).

### 3 Logical Spaces in *Scientific Representation*

Consider the roles that the notion of a logical space plays in *SR*.

In Part II of the book, one finds discussions of the problem of coordination (the process by which the formalism of a theory comes to have empirical content) and of measurement. The former is a self-conscious continuation of the conventionalist tradition of Poincaré, Reichenbach, and Grünbaum—although the problem in view is now taken to be fully general, rather than being limited to the special case of geometry. The account of measurement harks back to the discussion of logical space in van Fraassen's early publications.<sup>6</sup> These two discussions are tied together: “*The act of measurement is an act—performed in accordance with certain operational rules—of locating an object in a logical space*” (*SR* p. 165).

One of the highlights of Part III of *SR* is a discussion of a paradox of representation. A distinction is made early on in the book (p. 8) between phenomena (observable objects, events, and processes) and appearances (contents of observations or measurements). Several varieties of models are also distinguished (pp. 166 f.), including data models (compendia of appearances), surface models (processed and idealised data models) and theoretical models (the stock in trade of theories). A central claim of the book is that science represents phenomena as embeddable in theoretical models (pp. 238 and 247). This leads to a puzzle: while models can be identified with abstract objects, phenomena cannot be; but an embedding is a function—a map from one set of abstract objects to another. What sense, then, does it make to speak of phenomena as embeddable in a theoretical model? The answer, as I understand it, is a subtle one (see pp. 250–261). Let us suppose that I present a data model or a surface model as an adequate representation of some phenomena. This act should be thought of as locating myself in a logical space associated with my measurement concepts. The question whether the model I have presented is embeddable in a theoretical model associated with some theory is now a perfectly sensible one. But what about the gap between the concrete phenomena and the abstract representation of them that I have offered? *You* can say: “I see that the phenomena as represented by GB conform to the theory—but I worry

<sup>6</sup> See especially (van Fraassen 1967, 1970).

whether they really do.” But, on van Fraassen’s view, *I* cannot coherently say that, anymore than I can coherently say “It is raining but I do not believe that it is raining.”

There is much to worry over here. But let us stick to our course, and ask what, exactly, a logical space is. The notion can be found in many of van Fraassen’s writings. His favourite illustration is the sort of abstract spaces that are used to represent the possible colours of things.<sup>7</sup> We describe the colour of an object by locating it in such a space; each colour concept corresponds to a region of the space; and the structure of the space encodes meaning relations between our colour concepts (everything scarlet is red but not green etc.). In adopting a given colour space, we are not committing ourselves to the claim that each of its points corresponds to the colour of an actual object, but only to the weaker claim each coloured object can be assigned a location in this space.

In general, a logical space is an abstract space provided by a theory to represent the range of possible states or characteristics open to the objects described by the theory.<sup>8</sup> Typically a logical space will not just be a bare set but will come equipped with some sort of topological or geometric structure that tells us about relations of proximity or distance between the states or characteristics allowed by the theory. A region of a logical space corresponds to a possible property of an object, and the relations of inclusion between regions encode meaning relations between the corresponding predicates.<sup>9</sup>

What logical spaces we have available to us depends on what language we work in and on what theories we accept, and so will change over time (*SR*, p. 172). No surprise here: everyone will agree that with the advent of quantum mechanics physicists came to recognise the possibility of properties and states theretofore undreamt of, and also to recognise never before noticed relations of incompatibility between certain familiar properties.

But when combined with certain other themes of *SR*, this temporal dependence may seem unsettling. Near the end of his discussion of coordination, van Fraassen remarks that:

as Mach’s and Poincaré’s analyses show, measurement practice and theory evolve together in a thoroughly entangled way. Somewhat hesitantly one might say that the measured parameter [e.g., temperature or time]—or at the very least its concept—is *constituted* in the course of this historical development.<sup>10</sup>

And in his discussion of the paradox of representation van Fraassen seems at times to be headed towards a view on which, considered in abstraction from our descriptions of it, the world has no determinate structure. Certainly, van Fraassen argues vigorously against a popular alternative picture, on which associated with any concrete situation is a set of distinguished properties and relations, such that a perfectly accurate description of the situation could (in principle) be given by listing all facts obtaining in it concerning these distinguished properties and relations. And in the course of this argument he

---

<sup>7</sup> In addition to *SR* p. 164, see, e.g., (van Fraassen 1967, §4), (van Fraassen 1970, §III.4.b), and (van Fraassen 1980, §6.5.2).

<sup>8</sup> Aside from colour spaces, van Fraassen (*SR*, p. 164) mentions as examples of logical spaces “the PVT space in elementary gas theory, phase space in classical mechanics, Hilbert space in quantum mechanics; space and time themselves . . . .”

<sup>9</sup> On this strategy for handling meaning relations, see (van Fraassen 1967), (Stalnaker 1979), and (Arntzenius 1991).

<sup>10</sup> *SR*, p. 138. Van Fraassen hastens to add that the fact that temperature was constituted in this sense in the time of Galileo doesn’t prevent us from speaking coherently about facts concerning temperature at earlier times.

seems to deny (pp. 243 f.) that it makes sense to speak of the set of parts of an object in abstraction from our practices of representation; and he seems to assert (pp. 253 f.) that there is a close connection between the observation that any abstract domain can be equipped with many different relational structures and the thesis that a phenomenon, considered in itself, doesn't determine which structures are data models for it.

Remarks such as these may leave the reader wondering whether van Fraassen is ready to do more than just flirt with the wilder forms of neo-Kantianism—whether, for instance he is ready to join Kuhn in speaking of scientists separated by a scientific revolution as working in different worlds or to embrace Putnam's slogan that the world and the mind jointly make up the world and the mind. And, indeed, van Fraassen speaks not entirely disapprovingly of Kuhn's way of putting things and with measured approval of Putnam's.<sup>11</sup>

I think it is natural to feel some unease here. What sort of wild picture is being urged on us by this opponent of speculative metaphysics? But I suspect that van Fraassen will feel free to dismiss expressions of this unease so long as they are posed at a sufficiently global level—one of the themes of *SR* (see, e.g., p. 137) is that one can fall into incoherence in attempting to adopt a stance that prescind from too many of the presuppositions of our ordinary way of thinking of things.

So I will try to express my unease in a more local fashion, focussing on just one facet of our representation of the world, and discussing cases in which the agents in question share a great many presuppositions. In the next section, I will argue that in his early work at least, van Fraassen was committed to each of the following theses.

TI<sub>1</sub> Our cognitive situation plays a role in determining the geometric structure of the world.

TI<sub>2</sub> Relative to beings differently situated, our world might have a quite different geometric structure.

TI<sub>3</sub> Considered in abstraction from the cognitive situation of all beings, our world has no geometric structure.

Let us denote the conjunction of these three theses *TI*.

#### 4 Time as a Logical Space

Consider the discussion of what time is in §§III.4.b ff. of *An Introduction to the Philosophy of Time and Space*.<sup>12</sup> The point of departure is the contrast between Newton's view (time and instants are concrete entities) and Leibniz's (time is an ideal entity, the order of non-simultaneous events). A series of modifications and explications of Leibniz's view leads to van Fraassen's account of time as a logical space.<sup>13</sup> "We characterise the notion of *logical space* by saying that a logical space is a certain mathematical construct used to represent certain conceptual interconnections" (p. 104). On the view that van Fraassen extracts from Leibniz and Kant, time is a logical space with the structure of the real line,

<sup>11</sup> For discussion of Kuhn, see *SR* pp. 144 f. On Putnam's slogan, see (van Fraassen 2006, p. 125) and (van Fraassen 2009, p. 470 fn. 15).

<sup>12</sup> The first edition is (van Fraassen 1970). The second edition, (van Fraassen 1985), is unchanged except for the addition of a Postscript. All citations in this section are of this work. Citations by page numbers greater than 199 refer to the Postscript.

<sup>13</sup> For corresponding accounts of space and spacetime, see §§IV.3 and VI.6.c.

---

used to represent all possible temporal relations among events and the conceptual interconnections among these relations. (Thus, simultaneity is represented by identity of location on the real line, and the fact that temporal precedence is incompatible with simultaneity is reflected by the incompatibility of  $<$  and  $=$ .) (p. 102)

Implicit in the thesis that time *qua* logical space has a given structure is a commitment to certain temporal relations holding necessarily. But the notion of necessity involved is a subtle one because “the possibility that time has a beginning or that time is topologically closed cannot be ruled out, because . . . physical science might lead to such a conception of actual world structure” (p. 107). So “the actual temporal structure of events must necessarily be embeddable in our logical space. But the necessity here must be construed not as absolute logical necessity but as necessity relativized to the scientific theories that we accept” (p. 105). Van Fraassen appears willing to accept the counter-intuitive consequences of this view—e.g., that there is a sense in which there would be no time if there were no sapient beings (p. 107) and that there is a sense in which the structure of time may change if our theories change (p. 205).

Does this commit van Fraassen to TI? It looks like it does. Consider a world  $w$  containing agents  $A$  and  $B$  who agree about what sort of clocks should be used to measure temporal intervals and who both believe correctly that relative to these clocks, the total (past and future) temporal extent of matter at their world is a zillion years. But suppose that  $A$  and  $B$  disagree as to the structure of time at their world:  $A$  takes this structure to be given by the real line,  $B$  takes it to be given by a loop of length two zillion years. As I understand van Fraassen’s account, we should say that there is no fact of the matter about who is correct: relative to  $A$ ’s conceptual scheme, time is open and infinite; relative to  $B$ ’s it is closed and finite. Further, I don’t think that we can say anything much about the structure of time at  $w$  in abstraction from the conceptual scheme of any agent: it would seem that any feature that this structure has would have to be shared by any space suitable to be the logical space for time for some agent at  $w$ ; but in light of van Fraassen’s conventionalism concerning simultaneity (§V.3) as well as the temporal metric (§III.2), it seems unlikely that there is any such feature, except perhaps for a lower bound on cardinality. So it looks to me like van Fraassen’s account of time as a logical space commits him to each of TI<sub>1</sub>–TI<sub>3</sub>.

Now I suspect that van Fraassen would resist this characterisation. Already in the book on philosophy of space and time, there is an emphasis on how things look from ‘inside’ a philosophical view. Consider for instance, a case where we abandon the linear view of time in favour of the cyclic view of time. Would we then, on van Fraassen’s view, be committed to the thesis that time had changed its structure? No. For, in such a case “it is entirely accurate to say at each point in our history: time has such and such a structure, and that will not change. These assertions will be correct despite the fact that the structures referred to and the structural characters attributed are different at different times” (p. 205). In this way, van Fraassen may appear to avoid commitment to TI<sub>1</sub> and TI<sub>2</sub>. For it seems reasonable to think that if the structure of time depended on the cognitive situation of agents, then it would change when agents’ cognitive situation changed in a relevant way. But van Fraassen’s point in the passage just quoted is that from the perspective of agents who have through such cognitive change, the structure of time is invariant (they say that their earlier views were simply mistaken).

But I don't think that this gets van Fraassen all that he wants. For suppose that I too am an inhabitant of  $w$  and that and that I know that my view coincides either with the linear view of  $A$  or with the cyclic view of  $B$ , but I don't know which (perhaps they have been presented to me in some simpler form than I am used to and I am unable to see which corresponds to my own). I see that both views do justice to our shared principles of coordination and empirical beliefs. What should I say about the structure of time? It seems that I should say: "One of  $A$  or  $B$  is correct about the structure of time and one is wrong—and what determines which is which is which agrees with me rather than some fact about (the me-independent part of) the world." In this way, we recover the dependence of time on cognitive situation, even working within the first person perspective.

## 5 Four Questions for van Fraassen

1. *What distinguishes those parts of metaphysics that can contribute to understanding from those parts that are senseless?*
2. *Does the account of time as a logical space, as originally developed, imply TI?* Note that TI looks like it is a metaphysical thesis, according to the usual standards. (a) As noted above, Lewis's account of causation—which is a metaphysical thesis if anything is—commits him to the causal cousin of TI. (b) Some metaphysicians are sceptical that the contemporary debate concerning mereology is well-founded.<sup>14</sup> Attempts to explicate the sense in which claims about mereology are semantically defective lead to interesting metaphysical pictures which feature counterparts of the constituent theses of TI.<sup>15</sup>
3. *Does the account of time as a logical space qualify as metaphysical?* There are a couple of reasons for thinking that it should. (a) As just noted, it appears to imply the apparently metaphysical thesis TI. (b) The account was originally presented as arising via modifications of Leibniz's account, which it is natural to view as metaphysical since, e.g., it is in competition with Newton's clearly metaphysical account. The modifications in question don't look like it should affect this status.<sup>16</sup> And in any case, van Fraassen himself appears to view the logical space account of space, time, and spacetime as being in competition with the (paradigmatically metaphysical) absolutist approach.<sup>17</sup>
4. *How, if at all, does the picture of time as a logical space change when located in the more general account of SR? Do the answers from the previous two questions carry over?*

<sup>14</sup> Van Fraassen (2002, §1.5) shares their scepticism.

<sup>15</sup> See, e.g., the ontological anti-realism floated in (Chalmers 2009).

<sup>16</sup> One of the most important modifications is characterised as being Kantian: namely, the addition to the account of the claim that time has the same structure in any world. But this is *not* the sort of Kantian move that carries us out of the absolute-relational debate. For one thing, as van Fraassen notes, it is quite plausible that Leibniz himself already accepted the claim in question (van Fraassen 1985, p. 206).

<sup>17</sup> In addition to *SR*, p. 375 n. 12, see van Fraassen's (1972, p. 92) and (1985, p. 200).



---

**References**

- Arntzenius F (1991) State-spaces and meaning relations among predicates. *Topoi* 10:35–42
- Chalmers D (2009) Ontological anti-realism. In: Chalmers D, Manley D, Wasserman R (eds) *Metametaphysics: New Essays on the Foundations of Ontology*, Oxford University Press, Oxford, pp 77–129
- Jauernig A (2007) Must empiricism be a stance, and could it be one? How to be an empiricist and a philosopher at the same time. In: Monton B (ed) *Images of Empiricism: Essays on Science and Stances, with a Reply from Bas C. van Fraassen*, Oxford University Press, Oxford, pp 271–318
- Langton R (2004) Elusive knowledge of things in themselves. In: Jackson F, Priest G (eds) *Lewisian Themes: The Philosophy of David K. Lewis*, Oxford University Press, Oxford, pp 130–137
- Lewis D (1986a) Causation. In: *Philosophical Papers*, vol II, Oxford University Press, Oxford, pp 159–213
- Lewis D (1986b) Counterfactual dependence and time’s arrow. In: *Philosophical Papers*, vol II, Oxford University Press, Oxford, pp 32–66
- Lewis D (1986c) A subjectivist’s guide to objective chance. In: *Philosophical Papers*, vol II, Oxford University Press, Oxford, pp 83–132
- Lewis D (2009) Ramseyan humility. In: Braddon-Mitchell D, Nola R (eds) *Conceptual Analysis and Philosophical Naturalism*, MIT Press, Cambridge MA, pp 203–222
- Stalnaker R (1979) Anti-essentialism. *Midwest Studies in Philosophy* 4:343–355
- van Fraassen B (1967) Meaning relations among predicates. *Noûs* 1:161–179
- van Fraassen B (1970) *An Introduction to the Philosophy of Time and Space*, 1st edn. Random House, New York
- van Fraassen B (1972) Earman on the causal theory of time. *Synthese* 24:87–95
- van Fraassen B (1980) *The Scientific Image*. Oxford University Press, Oxford
- van Fraassen B (1985) *An Introduction to the Philosophy of Time and Space*, 2nd edn. Columbia University Press, New York
- van Fraassen B (1989) *Laws and Symmetry*. Oxford University Press, New York
- van Fraassen B (1991) *Quantum Mechanics: An Empiricist View*. Oxford University Press, Oxford
- van Fraassen B (2002) *The Empirical Stance*. Yale University Press, New Haven
- van Fraassen B (2006) One hundred and fifty years of philosophy. *Topoi* 25:123–127
- van Fraassen B (2007) From a view of science to a new empiricism. In: Monton B (ed) *Images of Empiricism: Essays on Science and Stances, with a Reply from Bas C. van Fraassen*, Oxford University Press, Oxford, pp 337–383
- van Fraassen B (2008) *Scientific Representation: Paradoxes of Perspective*. Oxford University Press, Oxford
- van Fraassen B (2009) Can empiricism leave realism behind? Toward a dialogue with transcendentalists. In: Bitbol M, Kerszberg P, Petitot J (eds) *Constituting Objectivity: Transcendental Perspectives on Modern Physics*, Springer–Verlag, Berlin, pp 459–479