

# Carnap's Second *Aufbau* and David Lewis's *Aufbau*

David J. Chalmers

I'm a longtime admirer of the Vienna Circle.<sup>1</sup> As a graduate student I was influenced by works by Moritz Schlick (*General Theory of Knowledge*) and Herbert Feigl (*The Mental and the Physical*). And Rudolf Carnap was long been one of my philosophical heroes. In recent years, almost everything I do has an element of Carnap in it. People sometimes find that surprising, because I hold metaphysical views about the mind that Carnap would certainly reject. But Carnap is such a rich philosopher that his ideas have power even for people who disagree with him about central issues.

I've been influenced by many works of Carnap, but the work that has gripped me the most is *Der logische Aufbau der Welt*—usually translated as *The Logical Structure of the World*. In the *Aufbau*, Carnap tried to construct the world, or at least construct descriptions of the world, from an extremely simple basis. I wrote a long book, *Constructing the World*, essentially trying to carry off a version of the project on somewhat different terms. I'm not talking about my own project here, though.

My main subject here is a second volume of the *Aufbau* that Carnap said he intended to write but never did. The first volume of the *Aufbau* attempted to construct the world from a phenomenal basis, or a basis of conscious experiences. Carnap said that the second volume would construct the world from a physical basis. I will sketch a picture of what the second *Aufbau* might have looked like, drawing on various works of Carnap's from around that period. I will argue that it would have been a philosophically viable work. Where the original *Aufbau* was widely dismissed as a noble failure, the second *Aufbau* could have been a philosophical success. Indeed, I will suggest that had it been written, it would have been a founding monument of 20th-century physicalism

---

<sup>1</sup>This paper is based on my Vienna Circle lecture on “The Carnap-Lewis Aufbau”, presented at the Institut Wiener Kreis at the University of Vienna on June 19, 2018. Thanks to Friedrich Stadler, Christoph Limbeck-Lilienau, and others at the Institute for helpful discussion. It's also based on part in a talk on “Lewis's Aufbau” to the David Lewis Society at the Pacific Division of the American Philosophical Association in 2011, and on a brief discussion of the Carnap-Lewis Aufbau on pp. 426-8 of *Constructing the World*.

and it would still be extremely influential today.

Another theme will be a deep connection between the work of Carnap and David Lewis. I will argue that Lewis's work can be seen as trying to carry off the project of Carnap's second *Aufbau*, and offers many valuable elements in doing so. Where Carnap's second *Aufbau* faces challenges, Lewis's work often offers a sophisticated solution. In this way I suggest that combining elements from Carnap and Lewis leads to a physicalist *Aufbau* that is still a living work today.

Let me back up with some history. Carnap published *Der logische Aufbau der Welt* in 1928, two years after he moved to Vienna. The official project of the book was to give a full description of the world in logical vocabulary alone, and then to construct all truths about the world from there. Most of the *Aufbau* focuses on constructing the world not from a logical basis, but from a phenomenal basis. Carnap takes a single primitive relation  $R$ , which can be understood as phenomenal similarity between two total experiences of a thinker. Using  $R$ , he first defines many other properties of experience, and then defines entities in the physical world. He constructs both low-level entities in physics and high-level entities in biology, psychology, and culture. Late in the *Aufbau*, Carnap dispenses with the relation  $R$  in an attempt to construct the world from a logical vocabulary alone.

For decades, the *Aufbau* has been viewed as a noble failure. The two most influential criticisms stem from its phenomenal basis. In *The Structure of Appearance*, Nelson Goodman argued that Carnap's construction of the qualities of experience from the relation  $R$  fails. In "Two Dogmas of Empiricism", Quine argues that Carnap's construction of spacetime points from properties of experience fails. These criticisms have been widely regarded as convincing.

A more general problem is that the *Aufbau* is widely regarded as committed to a form of phenomenalist reductionism about language and about reality: every expression can be wholly defined in terms of experience, and the entire world derives from experience. These phenomenalist theses have been almost universally rejected, and have led many to reject the *Aufbau* with them. In fact, Carnap did not have strong phenomenalist commitments. As various scholars of Carnap have pointed out in recent years, his strongest underlying commitment in the *Aufbau* is not to phenomenism but to a sort of structuralism, to serve the purposes of scientific objectivity.<sup>2</sup> Still, the association with phenomenism has stuck in the way that most philosophers think about the *Aufbau*.

It is well-known that in the *Aufbau*, Carnap says a number of times that the choice of a phe-

---

<sup>2</sup>For example, Friedman 1999 and Richardson 1998.

nomenal basis is somewhat arbitrary, and that he could have used different bases instead. On a number of occasions he mentions the possibility of using a physical basis. It is perhaps less well-known that Carnap in fact planned a second volume of the *Aufbau* that would be written using a physical basis.

Here is a note that Carnap wrote to Moritz Schlick in December 1927. Carnap lectured on the *Aufbau* in Vienna in 1925 and moved here in 1926. He circulated the manuscript and got a lot of feedback and by this point the *Aufbau* was reaching its final stages before its publication in 1928. But Carnap started to have second thoughts about the title of the book. His note was titled “Frage uber die Wahl des Buchtitels”, or “Question about the choice of book title”.

The title, “The Logical Structure of the World” appears to be in conflict with the work I plan for later. The constitution system of the book has an autopsychological (“solipsistic”) basis. At one point in the book, I briefly indicate that another constitution system with a physical (“materialistic”) basis is possible. The order is physical, psychological (without distinguishing the “I”), cultural. The result of this system is different. It serves not epistemology but science. Its basic domain is that domain whose processes uniquely possess the thoroughgoing univocal lawfulness. Mental and cultural objects are constituted by these physical objects.

Which of the two systems deserves more the name the “constitution of reality”? The first constitutes the world of knowledge or consciousness: perhaps one can also say: reality as known reality (one cannot talk about any others). The second system, however, can perhaps make a stronger claim to the name: it constitutes reality as the totality of everything that happens in space and time. Isn't it above all this viewpoint that we have in mind in science with the word “reality”?

I might choose the book title now already taking into account this later plan. Perhaps now “Logic of Knowledge”, and later “Logic of Reality”? Or now “The Logical Structure of Knowledge” and later “The Logical Structure of the World”.<sup>3</sup>

Although Carnap had written the first *Aufbau* intending to provide an objective basis for science, it seems that people including Neurath and Schlick were beginning to convince him that it was not quite as objective or intersubjective as he wanted it to be. Because the construction

---

<sup>3</sup>Translation taken from Pincock 2005 and Coffa 1991, both of which (especially Pincock) have useful discussions of the second *Aufbau*.

was starting from experience, it was serving epistemological purposes more deeply than it was serving the purposes of an objective description of reality. Starting in the late 1920s you find more and more remarks from Carnap saying that the *Aufbau* has a foundationalist epistemological flavor—many more such remarks than you find within the *Aufbau* itself.

Carnap suggests that this book with an experiential basis should be *The Logical Structure of Knowledge*, and the second one with a physical basis should be *The Logical Structure of the World*. In the end Schlick convinced him to keep the more ambitious name for the first volume. Sadly, Carnap never wrote the second *Aufbau* with the physical basis. Friedrich Stadler tells me that there is a lost manuscript called “part two of the *Aufbau*” that Carnap submitted for his habilitation. No one knows exactly what was in it. One thought is that maybe it could have had some elements of the second *Aufbau* with the physical basis. It’s certainly true that in Carnap’s physicalist period which followed over the coming years in the early 1930s where he wrote works on the unity of science, on psychology in physical language, and so on. You can certainly find clues there to what a second *Aufbau* might have been like. He just never turned it into a giant construction system the way he did with the first *Aufbau*.

This now gives rise to a speculative question. If Carnap had written the second *Aufbau*, what would it have been like? What could it have been like? Could one have carried off the project the intended project of the second *Aufbau* with a physical basis? If he had written it, how would twentieth-century philosophy have been different? My talk is going to be about these questions.

I should say that I am not a historian or a serious scholar of Carnap. These are merely some speculative remarks not well-grounded in scholarly attention to Carnap’s texts. But I think the speculative question is interesting. I’m going to sketch one speculative answer I’m going to be interested to hear from people who know much more about Carnap than I do about whether this speculation has any remote plausibility. If the historical speculation has none, as may well be possible, I think there is still a very interesting philosophical project in the neighborhood which is itself worthy of attention.

I will argue that Carnap could have written the second *Aufbau*, and that it would have been a successful work. I will also argue that David Lewis provides us with materials for a possible second *Aufbau* with a physical basis, one that is improved over Carnap’s second *Aufbau* in some respects. I will suggest that there is a hypothetical Carnap-Lewis *Aufbau* with elements from the work of each of them that might in some ways be the best of both worlds.

My discussion will have two main stages. I’ll talk about the basis on which this construction might work, the physical basis and the language for it. Then I’ll talk about the construction of the

rest of the world from this basis. I'll do both of these parts with reference to Carnap and Lewis. After that, I'll discuss the philosophical viability of the project and the historical role it might have played.

## The construction basis

Let me start with Carnap's construction basis. It's a structural or logical basis. Very early in the *Aufbau*, Carnap gives a statement about why structure is so important. He had this famous example of the railway map of Europe with different ways to characterize the railway map. We could talk about certain properties of certain railway stations and label them. This is the Vienna station, and this is the Prague station. You could just talk about relations, like all these ones are on the same line and have a relation for being on the same line. Or you could talk about the mathematical structure of the graph alone. These three sorts of sentence he calls property sentences, relation sentence, and structure sentences.

Carnap's idea is that we should turn property sentences and relation sentences into structure sentences. We start with property sentences with some one-place predicates. We try and get rid of our one-place predicates, like red, and structuralize those away with say, relations, like say, redder-than. So you replace all red claims by redder-than claims. This turns property sentences into relation sentences. Relation sentences still invoke multi-place relations, but ultimately you quantify those away. You say, there exists objects, some properties, and some relations such that the objects instantiate the properties and relations in such-and-such a way. This way you're left with structure sentences with a logical vocabulary alone.

The thought is that structure sentences yield maximal objectivity. No special undefined primitive properties or relations are used. The world sentence really should be a structure sentence.

In the *Aufbau*, Carnap's phenomenal basis initially involves relation sentences rather than structure sentences. He has a single undefined relation  $R$  of phenomenal similarity. The world-sentence looks something like

$$\exists x, y, z(Rxy \& Ryz \& \neg Rxz \& \dots)$$

The world-sentence says roughly that there are many things with this relation  $R$  holding between them in a certain complicated pattern. This is not yet a structure sentence. It's a relation sentence. Eventually, Carnap eliminates  $R$ , building a structure sentence. He does it in a very obvious flat-footed way, by turning the reference to  $R$  into a sort of quantified reference.

$\exists r \exists x, y, z (r(x, y) \& r(y, z) \& \neg r(x, z) \& \dots)$

This is a sort of application of what later came to be called the Ramsey-Carnap-Lewis method for defining and eliminating theoretical terms. Here we treat R as a theoretical term defined by its role in the world sentence, and we eliminate it. This leads to famous problems for structuralism. I'll get back to these issues shortly.

In sections 59 and 62 of the *Aufbau*, Carnap talks explicitly about the fact that he might have written the book with a physical basis. He gives us some helpful remarks about what the physical basis might be like. He actually gives three proposals for the physical basis. One of them involves starting with electrons and protons as basic objects and uses their patterns of acceleration or deceleration in space and time as a way of getting electromagnetism, gravitation, and then everything else. Another starts with relativistic worldlines for particles, and has basic relations for coincidence of worldlines and time ordering of points in a worldline. This construction derives from two earlier manuscripts where Carnap tries to construct spacetime from this worldline basis, although he doesn't try to construct all of physics or all of reality.

The most straightforward proposal for our purposes is his second proposal. In section 61 he says:

“We may choose as basic elements the space-time points of the four-dimensional space-time continuum and, as basic relations, their relative location in the continuum and the one-many relations between real numbers and space-time points which correspond to the individual components of the potential functions: the electromagnetic four-dimensional vector field and the tensor field of gravitation. According to the general theory of relativity in Weyl's formulation, all concepts of physical science can in principle be derived from these data.”

Carnap is proposing a certain basis here for his world-sentence. It's a description involving space-time points, the relative locations of those points, and the values of certain fields at those points. If we think about this picture metaphysically, it's a picture where in the world there are some space-time points, there are spatiotemporal relations between them<sup>4</sup> and there are electromagnetic and gravitational fields with values at those points. We can think of these fields as roughly corresponding to things like charge and mass.

---

<sup>3</sup>The manuscripts are “The logical foundations of kinematics” (1920) and “Topology of the spacetime world” (1924). See Howard 1996 and Pincock 2005 for discussion.

Now, Carnap doesn't really think about things in this metaphysical way. He uses the method of quasi analysis. His world sentence won't say explicitly that there are spacetime points. Instead it will just say that there are some things that stand in these relations. Then we can later define the notion of being a spacetime point in terms of those relations. His description of the world is that there exist some things that stand in these relations and have these field values.

This is already interestingly similar to a proposal from David Lewis. Lewis is famous for his Humean supervenience, the idea that everything in the world supervenes on what he calls the Humean manifold, a manifold of one little bit of the world after the other. In fact, it involves space-time points, one after the other with certain local qualities. Here's Lewis from "Humean Supervenience Debugged".

"Humean supervenience ... says that the fundamental relations are exactly the spatiotemporal relations: distance relations, both spacelike and timelike, and perhaps also occupancy relations between point-sized things and space-time points. Maybe those are needed. And it says that the fundamental properties are local qualities: perfectly natural intrinsic properties of points, or of point-sized occupants of points . Therefore it says that all else supervenes on the spatiotemporal arrangement of local qualities throughout all of history, past and present and future."

This quite reminiscent of Carnap's proposal in section 61 of the *Aufbau*, albeit with a more metaphysical flavor. Lewis's components are basically spacetime points, their spatiotemporal relations, and their local qualities. The points and the relations correspond directly to elements of Carnap's basis, and the local qualities correspond to Carnap's local values of fields at those points. Lewis also gestures toward an extra class of things, the occupiers of the points of the sphere, but it's clear that he would like to do without them if he could.

What are the primitive concepts used here? There will certainly be logical and mathematical concepts used. There will be expressions for spatiotemporal relations like distance relations in space or temporal relations, or spacetime relations. I'll simplify, ignoring relativistic considerations by assuming we have a spatial distance relation  $d_1$  and a temporal distance relation  $d_2$  between spacetime points. There will also be expressions for characterizing the gravitational and electromagnetic fields, or the local qualities like mass or charge. We'll see later on that these words can easily enough be Ramsified away. Lewis's description actually involves an appeal to perfectly natural qualities, which I'll get to shortly.

For now, Lewis's world sentence looks something like this.

$\exists x, y, z(d1(x, y) \& d2(y, z) \& m(x) \& c(y) \& \dots)$

That is, there exist some things that stand in these distance relations and have these masses and charges. Masses and charges are values of quantities rather than simple binary properties, but that's easy enough to handle. At this point, we still have quite a bit of primitive ideology. We've got  $d1$  and  $d2$ , we've got  $m$  and  $c$ . We're not yet close to logical structure.

Around here is where we use the Ramsey sentence method or as it's sometimes called the Ramsey-Carnap-Lewis method. This was a method in which all of these figures had some serious investment. The basic idea of the Ramsey-Carnap-Lewis method is that when a theory involving a theoretical predicate which is wholly defined by that theory, the theory is equivalent to a certain kind of quantified theory without the predicates. Say you have a certain theory about mass, let's say  $T(\text{mass})$ , which totally defines the theoretical concept of mass. Then the Ramsey method says that  $T(\text{mass})$  is identical to  $\exists m T(m)$ . That is, there exists some property  $m$  such that  $T$  is true of it. That will capture the entire content of your original theory of mass, without using the word 'mass'. These sentences will be analytically and a priori equivalent.

The Ramsey-Carnap-Lewis method is basically a way of turning a property or relation sentence into a structure sentence or at least into more of a structure sentence. We eliminate undefined expressions for properties like mass by quantifying over those properties and relations instead. If you do this repeatedly, you can replace all predicates and relations by logical expressions.

Carnap gets his name in this triple in virtue of work he did in the 1950s. In his papers on theories he really starts analyzing these sentences in depth and we have the famous Carnap sentences for certain conditional sentences that end up being a priori or analytic on this view. But if you look at the *Aufbau*, he's using a sort of Ramsification decades earlier.

The key use of Ramsification comes when Carnap dispenses with the relation  $R$  to produce a pure structure sentence as above. Carnap replaces  $\exists x, y(Rxy \& \dots)$  by  $\exists r \exists x, y(r(x, y) \& \dots)$ . This is a Ramsey sentence. The idea is that  $R$  is defined by its role in this sentence, so we just quantify and get rid of the reference to  $R$ . This Ramsey-style method is basically the natural method to go from an underlying characterization of the world in non logical terms to a characterization in logical terms.

Lewis uses the Ramsey method to dispense with mass and charge in his own way of doing things. He eliminates the reference to things like  $m$  and  $c$  by the Ramsey method of quantifying. For Lewis, the fundamental sentence doesn't say mass and doesn't say charge. It says something like:



$\exists$  (natural) properties  $m, c \exists x, y, z(d1(x, y) \& d2(y, z) \& m(x) \& c(y) \& \dots)$ .

Roughly, there exist some entities and some natural properties that have these structural features. We've moved closer to a structure sentence. We haven't moved all the way there though, as the spatiotemporal relations  $d1$  and  $d2$  are still present.

You might think we should do the same thing for space and time. Ramsifying space and time is a move I'm very sympathetic with. It sometimes goes by the name spatial functionalism, or spatiotemporal functionalism or spacetime functionalism. Just as we understand mass what plays the mass role in a theory and charge as what plays the charge role in a theory, we understand space and time as what plays the space and time roles in the theory, or spacetime as what plays the spacetime role.

One surprising thing about Lewis's philosophy is that it seems he never did this. You would have expected him to Ramsify space and time just as he does for charge and mass. Rather than taking  $d1$  and  $d2$  as primitive, he could have just said that there exist natural relations with their structural features. As far as I can tell, he never actually did this. I like to speculate that maybe if he had lived another few years, he would have Ramsified space and time as well. If we had, we would have a world sentence like this:

$\exists$ (natural) properties  $m, c \exists$ (natural) relations  $d1, d2 \exists x, y, z(d1(x, y) \& d2(y, z) \& m(x) \& c(y) \& \dots)$

That way you've got rid of all the primitive predicates, one place or multi-place, except perhaps for naturalness. You've got it down to logic and maybe a tiny bit more. This is a structure sentence very much in the spirit of Carnap's own structure sentence at the end of the *Aufbau*, except that it's a structural characterization of physics rather than experience. It's the sort of structure sentence that you'd expect for the physical basis in Carnap's second *Aufbau*.

Around here there's a famous problem, which is Newman's problem for structuralism. Any pure structure sentence is near-trivial as it is satisfied by any set with the right cardinality. This was pointed out by the mathematician Max Newman in 1928 in an article, "Mr. Russell's causal theory of perception", which critiqued Russell's *Analysis of Matter*. Newman points out that Russell's structuralism looks like it's near-trivial in just this way.

The same goes for Carnap's structure sentence:  $\exists r \exists x, y, z(r(x, y) \& \dots)$ . We can use this sentence to illustrate Newman's proof. The sentence is very easy to satisfy. If there's one set, say, of 200 objects that satisfies the sentence, then any set of 200 objects will satisfy that sentence. All you need to do is find a mapping of those 200 objects on to the original set and map them one to

one. Then find a relation  $r$  among the original 200 objects that witnesses the truth of the sentence, and map it to a corresponding relation  $r'$  among the new 200 objects. Then  $r'$  will witness the truth of the sentence too. At least if there are no restrictions on relations the sentence will be true for any set of 200 objects then any countably infinite set will make it true. This sentence at best constrains the cardinality of the world, and doesn't say anything to constrain the world beyond that.

It's well-known that Newman's problem was also discovered by Carnap around the very same year, in the *Aufbau*. Quite late in the *Aufbau*, in sections 153 through 155, Carnap suddenly acknowledges that he was supposed to find the logical structure of the world, but he still has the primitive relation  $r$  hanging around. So he needs to get rid of  $r$ . He comes up with the Ramsification method and moves to a structure sentence:  $\exists x, y \exists r(r(x, y) \& \dots)$ . Carnap is no fool and he immediately realizes that this has a problem. It is much too easily satisfied. He responds by requiring that the relation  $r$  be what he calls a natural or a founded or an experientiable relation. This requires that there is special class of relations which are natural. Then you get a sentence that looks like this:

$$\exists r \exists x, y (N(r) \& r(x, y) \& \dots)$$

This sentence is no longer trivial. But you might worry this is no longer the logical structure of the world, because the notion of naturalness isn't present in logic. Carnap says that we shouldn't worry, because naturalness is itself a basic concept of logic. Not too many people have been convinced by that claim, and Carnap appears to have been a bit embarrassed about it himself. He prefixes sections 153 through 155 by "Sections 153 to 155 may be omitted." (Nothing to see here, move right along.) That's very odd when this is the one part of the book that might vindicate the title *The Logical Structure of the World*.

Clearly Carnap realized there was a problem. You find a closely related move in Lewis at various points. Lewis also rediscovers Newman's problem in "Putnam's Paradox", as a problem for the related structuralist view he calls global descriptivism. He addresses it by an appeal to natural properties. He does things in a slightly different way. He doesn't build in naturalness to his descriptions, but instead takes it that all reference is constrained to be reference to natural properties and relations. He also doesn't claim that naturalness is a logical notion.

A more charitable way to understand Carnap's attitude toward naturalness is that whether or not naturalness is a logical notion, it's a broadly structural notion. An appeal to naturalness is perhaps more legitimate for an objective picture of the world than an appeal to mass and charge,

say. If we characterize the world by saying, there exists these entities and there exist these natural properties they have in such and such a way that at least it's a relatively structural description of the world, if not a logical description. There's no getting around Newman's problem. Everyone needs some non-logical notion in their base, and naturalness is arguably a way of retaining some kind of structuralism if not logicism here.

If you go this way, you're led to what we might call the Carnap-Lewis world sentence:

$$\exists m, c, d1, d2 \exists x, y, z (N(m) \& N(c) \& N(d1) \& N(d2) \& d1(x, y) \& d2(y, z) \& m(x) \& c(y) \& \dots)$$

Now we've gotten rid of  $d1$ ,  $d2$ ,  $M$  and  $C$  as primitive predicates. What's left? There's logic, mathematics, and naturalness. There's one primitive higher order predicate  $N$ , expressing naturalness of other predicates or properties. In principle one can Ramsify any physical theory this way. You might find something fairly explicitly like this in Lewis, except that he doesn't Ramsify space and time.

It's natural to think that if Carnap had written his second *Aufbau*, it would have gone something like this. Carnap gives us clues about what his physical basis is going to look like right there in section 62 of the *Aufbau*. He also gives us very strong clues as to what his preferred structuralization method is. It's Ramsification along with an appeal to naturalness. If you combine his suggested physical basis in section 62 with his preferred structuralization method in sections 151-53, you will very plausibly end up with something like this. It is very close to what Lewis ends up with, too.

One might worry about whether even after invoking naturalness, a Ramsey-style structure sentence along these lines still undetermines reality. After all, it says that there are certain natural properties but does not specify what they are. In "Ramseyan Humility", Lewis argues that natural properties such as mass and charge can be replaced by distinct alien natural properties in other possible worlds. As a result, a purely structural base doesn't form a supervenience base for all truths, and a metaphysical supervenience base needs to be a non-Ramsified version which names mass and charge as distinct properties. However, on Lewis's view the relevant metaphysical distinction between natural properties (in terms of "quiddities") turns out to be inexpressible in language and ungraspable in thought. As a result, the Ramsified version still serves as an epistemological basis from which all expressible truths can be derived. For Carnapian purposes this is probably good enough. Furthermore, it seems unlikely that Carnap would have accepted quiddities underdetermined by structure, so I set the issue aside for our purposes.

For the part of the project concerned with the physical basis, it is arguable that Carnap doesn't need that much help from Lewis. Carnap could easily have come up with a world-sentence along these lines himself. If he took his preferred descriptions of the world in terms of physics, and applied his preferred method of Ramsifying and invoking naturalness, he would have ended up with something along these lines. This gives us at least a possible picture of the basis for Carnap's second *Aufbau*.

## The construction method

The next step is the construction of everything else from the physical basis. A construction system requires a method of construction (what Carnap calls an "ascension form") to construct new entities from entities in the basis. Putting things linguistically, we start with a world-sentence using a limited class of expressions and we have principles for deriving new sentences from there, ultimately encompassing all truths about the world.

Carnap's central construction method is to use definitions. One defines new expressions in terms of old expressions. Using definition sentences, we can derive truths in a broader vocabulary. For example, if we already have  $\exists x(\text{male}(x) \& \neg \text{married}(x))$  in our system, and if we also have the definition  $\forall x : \text{bachelor}(x) \text{ iff } (\text{male}(x) \& \neg \text{married}(x))$ , then we can derive  $\exists x \text{bachelor}(x)$ .

Definitions are usually required to meet criteria of adequacy. Standardly definitions are required to be analytic, a priori, or necessary (often all three). Carnap's criteria for adequacy in the *Aufbau* are notoriously problematic. He requires explicit definitions, which is a very strong requirement that he doesn't always follow. At the same time, there's no requirement that a definition be necessary or analytic or a priori. He simply gives the very weak requirement that a definition be extensionally adequate. By these lights we could say that "Someone is a bachelor iff they are an untidy man" is a definition, as long as it happens the two classes coincide.

In the preface of the second edition of the *Aufbau* Carnap says that he got these things wrong. First, you can't do everything with explicit definitions. You're going to need sometimes implicit or contextual definitions. Second, to serve his purposes, it's not enough that the definitions merely be extensionally adequate. He says the extensiveness must not be accidental, but necessary. It appears that the definitions must be necessarily true. For Carnap necessity and analyticity were very closely connected so there maybe this was close to requiring that the definitions be analytically true, or true in virtue of meaning.

Something like this is very plausibly required if you want this kind of project to serve con-

ceptual and epistemological purposes as Carnap did. It's also required for a world-sentence to be considered a complete descriptions of the world in any reasonable sense. In Lewis-style constructions you always find criteria like that imposed. Even in Carnap's original *Aufbau*, it often looks as if he is giving definitions that meet the stronger standards.

In thinking about Carnap's second *Aufbau*, I will take Carnap's remarks here at face value and assume that the construction system works through definitions that are something like necessary, a priori, and analytic. We'll start from a physical basis, and define further expressions in terms that can be derived from the physical bases.

How exactly will derivation from a physical basis work? In the original *Aufbau* Carnap works with an experiential basis and spends most of his time deriving truths about experience and about the physical world from there. He only very briefly discusses the derivation of other things from a physical basis. He discusses other minds, culture, and values, but this all goes by extremely fast.

The best guidance for how Carnap's construction from a physical basis might go comes in two classic works from 1932, at the height of Carnap's physicalist period: "Physics as a Universal Language for Science" (published as a book, *The Unity of Science*, in English in 1934) and "Psychology in Physical Language". The first of these works argues that all scientific claims can be translated into the language of physics, roughly by defining high-level scientific notions in observational and physical terms. The second gives an analysis of mental or psychological claims (including observational claims) in physical and behavioral terms, so that they can be derived from a physical description of the world. Presumably Carnap would have used methods like these for construction from a physical basis, if he had been writing a second *Aufbau* during this period.

These works were not a second *Aufbau*, of course, and Carnap's remarks about construction are sometimes sketchy and unsatisfying. His definitions also sometimes turn on behaviorist and phenomenalist claims that would now be widely rejected. Around this point, we can get some help from the work of David Lewis. In a way, Lewis's life work involved the construction of the world from something like a physical basis using conceptual analyses. Because Lewis was working decades later, his analyses are often much more sophisticated than Carnap's, but they share a similar spirit. Triangulating between Carnap and Lewis, we can get a sense of what a successful construction project in the vicinity of Carnap's second *Aufbau* might look like.

Lewis's official project is reduction of all elements of reality via Humean supervenience, or supervenience on his physical bases as described earlier. For Lewis, supervenience connections are always supposed to be necessary, and they were grounded in some sort of conceptual analysis. That's an approach often associated with followers of the Canberra plan like Frank Jackson, and

it's also central in Lewis's system. This kind of reduction by conceptual analysis is well-suited for our purposes, because it provides the kind of construction that might serve the goals of the second *Aufbau* with construction going through definitional connections.

How will the construction work for Carnap and for Lewis? We start from the structural world-sentence. Let's say we have the fully Ramsified Humean world-sentence goes a step beyond Lewis by Ramsifying space and time. The first thing we need to do from there is recover the non-Ramsified world-sentence with space-time, mass, charge and so on. You can do that using definitions of spatiotemporal relations, mass, and charge in the style of Carnap sentences. Mass is defined is what plays the mass role, and similar for charge and distance relations.

What exactly are the mass role and the distance role and so on for the purposes of this derivation? One way to do cash this out is by their role within the basic world-sentence itself. The trouble is that it isn't plausibly analytic or a priori that mass plays its full role in the world-sentence. There will be a lot of empirical and contingent details there about the distribution of mass that don't look at all as if they are part of the concept. Better may be the role within a certain core scientific theory of the world: for example, maybe mass is at least roughly what plays a certain inertial role in Newtonian physics, or distance may be what plays a certain role in relativity. Something like that will arguably be a priori, especially if the relevant notions are taken to be regimented theoretical notions like Newtonian mass or Einsteinian distance. This will enable us to recover distance, mass, and so on at least as they play a role in our fundamental theories. All this is very much something that both Carnap and Lewis could have done. This gets us to the world of physics, with spacetime points and their basic physical properties.

How do we construct nonfundamental objects, from molecules and cells through tables and tigers to stars and galaxies? There are two issues here: ensuring the objects themselves exist, and ensuring they satisfy the right predicates. Let us first focus on the construction of the objects.

Carnap's remarks about constructing high-level objects are sketchy. In general, he regards them as collections of lower-level objects. He first constructs "visual things" as collections of world-points which have previously been constructed from visual experience. He later constructs "perceptual things" more generally and identifies some of these as physical things. A human body, for example, is a visual thing and a physical thing. In any case, it seems clear that objects are identified with collections of points.

Lewis understands high-level objects in a closely related way: as mereological fusions of basic objects. Assuming that Lewis's basic objects are spacetime points (and not their occupants), this is close to Carnap except in using the mereological concept of fusion rather than the set-theoretic

concept of collection. Fusions are defined in terms of parts: A is a fusion of some things when (i) those things are all parts of A and (ii) if those things are part of some object B, A is a part of B. Lewis then gets high-level objects from basic objects by the principle of universal composition: whenever some things exist, a fusion of those things exists.

Lewis took universal composition to be an analytic principle. It's controversial among metaphysicians whether that's an analytic claim. Many deflationary metaphysicians agree that it is, while many inflationary metaphysicians deny it. Of course Carnap was a deflationary metaphysician, and his later views in *Empiricism, Semantics, and Ontology* are clear here. He has no problem with a claim of this sort being analytic, at least within a framework. Presumably within the universal mereology framework, it is analytic that any things have a fusion. In the set-theoretic framework things are more complex: if one says it is analytic that any things form a collection, one is threatened by Russell's paradox. Perhaps a more restricted claim will serve Carnap's purposes: for example, that any things which are not collections form a collection. This is perhaps one reason for preferring the mereological construction of high-level objects to the set-theoretic construction.

In any case, for Carnap this will certainly be a pragmatic choice of framework. As always, we shouldn't get too worried about whether this framework is the true framework for reality. That's a meaningless external question for Carnap. What matters is that the framework is useful. All this is explicit in ESO, but you find this sort of talk about metaphysical frameworks already in works like the *Aufbau*. It seems he'd be happy with either mereological or set-theoretic frameworks for the construction of high-level objects.

One interesting question is whether the appeal to mereology requires the notion of parthood as basic ideology. I don't think Lewis ever really defines it away. The same goes for Carnap's notion of collection. Even if these mereological or set-theoretic notions are required, they are arguably broadly structural notions at least in the way in which naturalness is a structural notion. So they are consistent with this broadly structural project.

What about construction of high level predicates, like being a cell or being a bottle, or being a computer? Around here things are always sketchy. No philosopher has tried to give analyses for every high-level predicate, or even to give truly satisfying analyses for more than a very few. Still, we can at least ascertain a general approach.

Carnap spends little time on high-level predicates in the *Aufbau*, with the exception of a few special cases. He gives a five-part analysis of "my body" in section 129, and sketchy analyses of "organism" in section 137 and certain cultural objects in sections 151-2. He goes into more detail about these matters in section 5 of "Physics as a Universal Language for Science", con-

centrating on expressions used in high-level sciences. He says that that there is no question that physical language is applicable to the inorganic sciences (chemistry, geology, astronomy), and that terminology here can always be defined “in terms either of physical determinations or qualitative determinations (e.g. results of observations)”. Biology is more complex because the question of vitalism makes it controversial whether biological laws can be reduced to physical laws. Even so, biological concepts (species, organisms, organ) can be defined in terms of physical concepts. He gives a rough and ready definition of ‘fertilization’ in terms of union of sperm and egg, where these are defined as cells of specified origin and specified perceptible properties. In general, his view seems to be that we can define all the key expressions in the natural sciences in physical and observational terms.

The view that all expressions in the natural sciences can be defined in physical and observational terms is now widely rejected. Theoretical terms often do not have the straightforward connection to physics and observation that Carnap’s approach requires, and many scientific kinds are multiply realizable at the physical level. Still, a descendant of this view associated with Lewis and others remains popular. This is the approach in which one defines scientific terms in broad structural or functional terms, so that various properties are defined in terms of a structural network of interaction with each other and with observation. Lewis typically defines high-level predicates in functional and observational terms along these lines. Some high-level predicates are defined in terms of other high-level predicates, but the hope is that if we define them all at the same time, the entire network will be derivable from underlying truths about physical properties and observational properties?

What about construction of observational properties? These are properties we seem to directly observe in the world, like colors. Maybe shapes and spatial properties fall here, though it’s tricky because those are already present in the base. Whether the observational notion of space is the same as the one in the base is itself a substantive question. Earlier we suggested using a theoretical notion of space in the base, in which case the observational notion may be distinct. Carnap in *Der Raum* distinguishes multiple notions of space (formal space, physical space, intuitive space) and might be inclined to do that here. It’s an interesting question what Lewis would make of a similar move.

In the *Aufbau*, Carnap defines observational predicates such as colors in terms of the mental states they are associated with. In “Physics as the Universal Language for Science”, his basis observational truths are protocol sentences about what people experience at certain times: for example “thirst now” and “red now (seen by S)”. This has the same effect. Observational claims are



reduced to claims about mental states. Lewis takes a more complex version of the same strategy, defining observational predicates like colors in terms of causal/nomic relations with mental states. Red things are what normally bring about a certain kind of experience. He interdefines the color properties and color experience properties in a complicated way in his paper “Naming the Colors”. He didn’t do the same for space, but there’s an analogous project for defining observational space in terms of relations to spatial experiences. That way observational space is defined differently from physical space (as they are for Carnap in *Der Raum*), but they may still end up roughly coextensive.

This leaves us with the construction of the mental. In “Physical Language as the Universal Language for Science”, Carnap argues that protocol sentences such as “thirst now” will be deducible from full enough descriptions of the nervous system once we know enough about neuroscience. After worry that such a deduction is “utopian”, he points to the more straightforward project of defining these in behavioral terms. For example, ‘seeing red’ is defined partly in terms of saying ‘red’ in response to the stimulus ‘What are you seeing now?’, and in terms of pointing to certain colors on a card when asked, and so on. He develops this approach at far more length in “Psychology in Physical Language”, giving detailed behaviorist treatment of various mental claims and arguing these can be derived from the physical level.

Of course behaviorism is now almost universally rejected, and behaviorist analyses like Carnap’s have principled problems. For example, couldn’t someone see red even with no capacity to behaviorally respond? But once again, a more sophisticated descendant of the view can be found in the work of Lewis and others such as David Armstrong. Lewis and Armstrong famously define mental states functionally, in terms of their interaction with each other as well as with inputs and outputs. These functional definitions are in the spirit of behavioral definitions but go a step beyond them by including causal connections between internal states in the definition. This removes many of the most serious problems for behaviorist analyses. For example, a paralyzed person can still see red in virtue of having a state with the right internal role.

Lewis gives broadly functional analyses of mental notions like experiences, beliefs, and so on, in terms of causal relations to each other and to behavior. As always, the project is mostly schematic. For some notions he goes into a bit more depth than others. But if we understand them all as functional properties, then the idea is that truths about them will be derivable from truths about patterns of causation which are ultimately grounded in the physical base. Of course like any philosophical view this functionalist view is controversial, but it’s been very influential and is still one of the leading views about mental states.

There are also more complex human domains such as culture, values, and morality. Carnap discusses culture and values in a famously swift part of the *Aufbau* (sections 150 through 152), giving a behavioristic analysis of culture and a mentalistic analysis of value in terms of value experiences. (Earlier, in section 59, he also unexpectedly discusses a possible derivation of value from energetics.) He says very little about ethics per se in the *Aufbau*, but he makes clear at the start of “Physical Language” that he rejects the questions of ethics as metaphysical (though human moral psychology may be an empirical science). Lewis addresses value and morality at more length in a couple of papers. His rough idea is that something is good if it is what we would desire to desire, and an act is right if it produces the most good. Of course all this is controversial, but again it’s a philosophically viable project.

As for construction of values, this goes by extremely fast in the *Aufbau*. Actually I quite like the part where, I was intrigued by this part of the *Aufbau* where Carnap says he thinks he can reconstruct values from energetics. Someone’s theory of energetics are going basically give you values in it. I haven’t been back and looked at the relevant theory of energetics. Carnap was not the most, he didn’t have the most subtle view of the evaluative and the cultural, but I think that can be forgiven. Lewis’s view also probably had some problems, but again everyone has issues with the evaluative, and it’s not the kind that makes a project a noble failure.

There are some domains Lewis addresses at length that Carnap says relatively little about. These include the nomic domain (truths about laws of nature), the causal domain (truths about what causes what), and the modal domains (truths about what is necessary and possible). Carnap addresses laws and causation very briefly in section 165 of the *Aufbau*, giving a simple regularity-based analysis of each: roughly a law is a regular connection between qualities at different locations. He distinguishes different sorts of law in terms of the spatial and temporal relations between the relata: state laws (simultaneous relata), process laws (one follows the other), proximity laws (one close to the other), causal laws (process law with temporal proximity). Cause and effect are defined as relata of a causal law.

There are well-known problems for these simple analyses of laws and causation. Again, Lewis gives more sophisticated analyses in a similar spirit. For causation, Lewis constructs laws of nature from underlying Humean basis via a best-system analysis. Laws are the regularities that appear in the best system for summarizing the regularities, where the best system is defined in terms of strength and simplicity. Then he goes on to construct counterfactuals from laws and modal truths, he constructs causation from counterfactuals, and this yields a whole realm of broadly causal or nomic notions.

Around this point, Lewis takes a notorious turn into the modal realm. He brings in claims about necessity and possibility to analyze many of these causal and counterfactual notions. And he grounds these claims using his modal realism with concrete possible worlds. You might wonder how Lewis' modal realism fits into a project in the spirit of Carnap's second *Aufbau*. The rough answer is that it doesn't. I propose to subtract that part of Lewis from the system. You don't need Lewis' modal realism for this project, and the project is much more plausible without it, because modal realism is notoriously implausible. Certainly Lewis's modal realism would horrify Carnap.

There's an interesting question about how modal truths themselves will be derived for Carnap. He doesn't talk much about modal truths in the *Aufbau*, but twenty years later in *Meaning and Necessity* these are his main focus. I think the analyses of necessity used there could certainly be used to derive modal truths that could help out in the Lewisian derivation of counterfactuals and causation. Carnap roughly says that a sentence is necessary when it is true by definition, so it looks like we don't need too much more than definition to get these modal truths by his lights.

## **Is the physicalist *Aufbau* philosophically viable?**

So far I've made the case that a physicalist *Aufbau* was possible, and I've fleshed out two pictures of how it might have gone, deriving from Carnap and Lewis respectively. Carnap's system is sketchier and more simplistic in various ways, while Lewis's is more detailed and more sophisticated. Nevertheless, they share a spirit. We might say there are at least two related *Aufbaus* here.

First, there is Carnap's second *Aufbau*. This is the *Aufbau* he might have written using ideas from the first *Aufbau* and from his physicalist works of 1932. It starts with a structural basis of elements (which we know as spacetime points) with founded relations between them and values of fields at them. It constructs physical predicates from there using a Ramsey-style definition. It constructs causation and laws from regularities. It constructs high-level objects by set-theoretic collection. It constructs high-level physical predicates by observational and physical definitions. It constructs psychological predicates by behavioral definition.

Second, there is Lewis's *Aufbau*. This is an *Aufbau* that Lewis in effect wrote in multiple parts over the course of many different works between the 1960s and the 1990s. It starts with the Humean supervenience basis of points with spatiotemporal relations between them and natural properties that hold at these points. It identifies mass and charge from there using Ramsey-style definitions. It constructs laws from best-system regularities, counterfactuals from laws and possi-

ble worlds, and causation from counterfactuals. It constructs high-level objects by mereological fusion. It constructs high-level physical predicates using Ramsey-style structural/functional and observational definition. It constructs psychological predicates by functional definition.

Are these *Aufbau* projects philosophically viable? I'll start with Carnap's second *Aufbau*. Certainly it would be subject to serious objections, as any philosophical project is. But would it be subject to the sort of fatal flaws that made people regard Carnap's actual *Aufbau* as a noble failure?

We can start with some of the most famous problems from the first *Aufbau*. I think it's clear that the Carnap's second *Aufbau* doesn't suffer from Goodman's or Quine's problems for the phenomenal basis, since it doesn't use a phenomenal basis. It is subject to Newman's problem for a purely logical world-sentence. There's no getting around that. Because of Newman's problem, any construction system needs something extra-logical in the base. The second *Aufbau* I've spelled out uses naturalness in its base. But this is reasonably taken to be a structural notions, if not a logical notion. The second *Aufbau* might not have yielded the logical structure of the world, but it still might in some sense yield the structure of the world. And the resulting structure could perhaps still be seen to be objective in a way that suited Carnap's purposes.

Carnap's second *Aufbau* as I understand it involves analyticity and conceptual analysis, which are a major part of Quine's critique in "Two Dogmas". This part of the critique would certainly apply. My own view is that Quine's criticisms are not terribly strong. Carnap actually had the stronger part of this debate. I've recently argued that ideas from Carnap's "Meaning and synonymy in natural languages" can be extended to rebut Quine's arguments. I think these days many people agree that Quine's critique is not as strong as was once thought.<sup>5</sup> So I don't think there is a fatal objection to Carnap's second *Aufbau* here.

Of course some people will still object to the appeal to analyticity and apriority. And some people who accept these notions will be doubtful whether one can find analytic or a priori entailments from physical truths to other truths as Carnap, Lewis, and others think. But again, this is just ordinary philosophical disagreement and a matter of live philosophical debate.

A more serious objection to the second *Aufbau* is that it requires the use of definitions. We need definitions of all non-physical expressions in order to derive all truths from our physical basis. And today, many people deny that such definitions are available. Almost every definition anyone has ever proposed seems subject to counterexamples.

I talk a lot about this in my book, *Constructing the World*. I argue that instead of definitions we

---

<sup>5</sup>See Bourget and Chalmers 2014: 70PhilPapers survey accept or lean toward accepting the analytic/synthetic distinction.

should use the weaker relation I call scrutability, so that high-level truths can be a priori entailed by base truths. And I argue that even if there are counterexamples to every definition, this is no obstacle to scrutability. That's how the project would go given by own commitments. But for present purposes I'm trying not to bring in my own commitments and the apparatus I like. The *Aufbau* project as Carnap conceives it uses definitions. How far can we get on those terms?

The problem is that all definitions seem to have counterexamples. I think this is not as big a problem for Carnap as for some others, because for Carnap the project of defining terms is not conceptual analysis but conceptual explication. In explication, we're allowed to replace our concept by new and revised concepts. They're changed a bit around the edges. They're more precise in certain ways and better behaved. When a proposed definition has an intuitive counterexample, that's usually a fatal problem for a conceptual analysis, but not for an explication. We can say that it's a revisionary explication. In effect we're not analyzing tables, but we're analyzing Tables. As a result, we construct the truth about Tables, which is good enough for Carnap's purposes.

So, I think that the most well-known fatal problems for Carnap's first *Aufbau*—Goodman's and Quine's critique of the phenomenal basis and Newman's problem for the logical basis—are not problems for the second *Aufbau*, at least appropriately understood. The problems of analyticity and definitions still apply, but these are far from clearly fatal, especially if we allow that definitions can be revisionary in the spirit of explication.

Are there other, new, fatal problems for Carnap's second *Aufbau*? Certainly there are problems that could be raised. The most obvious worries are his reliance on simplistic analyses: observational analyses of high-level predicates, behavioral analyses of mental predicates, simple regularity analyses of laws and causation, and so on. In every case, the sort of analysis that Carnap gives are now regarded as inadequate. So the second *Aufbau* that Carnap would have written in the late 1930s could not be quite right by present lights.

Still, in each of these cases, Lewis's work provides a more sophisticated treatment in the same spirit that avoids the worst problems. For high-level predicates he gives structural/functional analyses, for mental states he gives functional analyses, for laws he gives a best-system analysis, for causation he gives a counterfactual analysis. Lewis's general strategy is to provide a holistic analysis where Carnap gives an atomistic analysis. Laws are analyzed in terms of a whole system of laws, mental states are analyzed in terms of a whole system of mental states, theoretical terms are analyzed in terms of a whole system of theoretical entities and theories, counterfactuals are analyzed in terms of a system of worlds. The holistic analysis is still amenable to grounding in the physical, while avoiding most of the problems for atomistic analyses.

Looking at Carnap's second *Aufbau* this way, its problems are not fatal flaws. They are merely inadequacies to be rectified by better holistic analyses in much the same spirit. Lewis's work provides the necessary rectification. When we look at Carnap's second *Aufbau* in light of the work of Lewis and others, it is not a noble failure at all. Instead it is the start of a thriving research program.

It is easy to imagine an improved version of the second *Aufbau* with elements from both Carnap and Lewis. I think of this as the Carnap-Lewis *Aufbau*. It is intermediate between Carnap's second *Aufbau* and Lewis's *Aufbau*, with elements from both. It is roughly what one would get if one took Carnap's second *Aufbau* as described above, and replaced Carnap's simple analyses of high-level predicates, mental states, laws, and causation with Lewis's more sophisticated holistic analyses. It is also what one would get if one took the core elements of Lewis's system, Ramsified space and time so that the basis is structural, and eliminated the modal realism. This Carnap-Lewis *Aufbau* would arguably have the best of both worlds.

Does the Carnap-Lewis *Aufbau* have fatal problems? Certainly there will be philosophical objections. Many nonHumeans (including me) would reject the construction of laws and causation from regularities in a Humean basis. But this is just regular philosophical disagreement, and many are on the Humean side here.

Similarly, many would reject Lewis's construction of the mental from the physical. I certainly would. My own view is that you cannot deduce facts about consciousness from physical facts. In fact, I am inclined to think that Frank Jackson's knowledge argument is a refutation of any structuralist *Aufbau*-style project. Mary in the black and white room could learn any structural facts (since the whole point of structural facts is that they're objective and can be understood by anyone), but would not be in a position to know key facts about the experience of color. So I think any structuralist project fails as an account of all of reality. Still, again this is standard philosophical disagreement, and Lewis's view is philosophically respectable. Every view on the mind-body problem has problems, and Lewis's is one of the more viable. So the problems here are not the sort of fatal flaw that renders the project an obvious failure in the way that Goodman's and Quine's criticisms of the first *Aufbau* made it an obvious failure. The issue here is more a matter of one's philosophical commitments.

There are perhaps some potential issues for indexical or *de se* truths, such as 'I am in Vienna' or 'It is 5pm now'. It's hard to see how one can derive them directly from a physical basis without indexicals. Carnap's way of handling this issue was to try to derive objective counterparts of these truths, e.g. 'Such and such person is in Vienna at such-and-such time'. Perhaps that makes sense if

he's trying only to handle objective reality. Lewis is clear that he thinks *de se* truths like this cannot be derived from more objective truths. He doesn't build them into his metaphysical basis, but he needs them for a full epistemological basis. In my own work I've also included indexical elements in an epistemological basis. Of course this moves beyond a purely objective basis since *de se* truths are paradigmatically subjective. It's an interesting question to what extent this phenomenon undercuts the whole *Aufbau* project of deriving everything from an objective basis. Perhaps one could make the case that a small augmentation doesn't undercut the project, or one could restrict the scope of the project to objective truths as Carnap wants to. Either way one will be left with a modified project at least in the style of the *Aufbau*.

Another minor issue is whether one needs a special that's-all element in the basis to handle negative facts, like "There are no angels". It looks like positive physical facts alone don't entail that there is nothing nonphysical, so one needs a further claim along the lines of "Everything is physical" or "There's nothing fundamental I haven't mentioned". I don't know whether Carnap or Lewis address that explicitly. In my work I've augmented the physical basis with a that's-all element as well as phenomenal and indexical elements, yielding a world-sentence I call PQTI (physical, qualia, that's-all, indexicals).

Other problems? Of course there are questions about how this sort of project stands in light of quantum mechanics. In quantum mechanics, it looks like the base will involve a wave function. This may mean rejecting Carnap's and Lewis's proposed base of spacetime points alone. The wavefunction involves a sort of entanglement where properties of spacetime points are no longer independent of each other. This may mean rejecting or modifying Lewis's Humean supervenience, as Lewis himself acknowledges, but it is not a major obstacle to Carnap's physicalist project generally. We simply need to move to a new base reflecting this new physics, with a structurally characterized quantum wave function in the base. There are also notorious problems with deriving the macroscopic world from the quantum wave function. I don't think either Carnap or Lewis had solutions to those problems. Still, I argue in *Constructing the World* that the problems are solvable in principle, though many of the details will depend on the interpretation of quantum mechanics you embrace. In any case, quantum mechanics is a serious problem for everybody, so there is not a unique problem for the *Aufbau* project here.

So where does the Carnap-Lewis *Aufbau* stand, philosophically? I think that although there are challenges for this project, as there are for any project, there's no problems like the Goodman/Quine problems that render the project an obvious failure. Carnap's version of the project has some weaknesses that can be fixed by moving to Lewis's more sophisticated holistic analy-

ses. Lewis's analyses are subject to the usual slew of philosophical objections here and there, but his views (with the possible exception of modal realism) are still philosophically alive today and philosophically viable.

In a similar way, the Carnap-Lewis *Aufbau* is a living work. It's certainly within the space of live philosophical views. The Carnap-Lewis *Aufbau* is not an obvious failure. It is a still viable philosophical blueprint for the world.

## Historical questions

Historically: could Carnap have written his second *Aufbau*? I don't see why not. Most of the needed elements are present at least in a rudimentary way either in the *Aufbau* or in the physicalist works of 1932. The shape of the structural physicalist base is strongly suggested by elements in the first *Aufbau* alone. The construction method for high-level and mental predicates is spelled out in the 1932 works. Other details can be found here and there in each.

A stronger speculation is that if Carnap had set out to write a second physicalist *Aufbau* in the late 1920s or early 1930s, he *would* have written something like the work I describe. Of course this is a historical counterfactual, and I know historians usually don't like counterfactuals. I'm not a historian, though, so I can speculate. I think that if he had written a second physicalist *Aufbau* with a natural combination of his views from the 1928 and 1932 works, something like the second *Aufbau* I have describe would have been the result.

Could Carnap have written the Carnap-Lewis *Aufbau*? Or at least, would he have endorsed it? Here we are on more speculative territory. Lewis's ideas here date from the late 1960s and early 1970s, forty years on. It's unlikely that Carnap could easily have jumped those decades and come up with Lewis's analyses just like that. Still, he wasn't so far away. Especially the Carnap of the 1950s has full use of the Ramsey-Carnap-Lewis method for more sophisticated functional analyses of many of these notions. So perhaps if the Carnap of the 1950s had been interested in writing a physicalist *Aufbau*, it would have had more elements of the Carnap-Lewis *Aufbau*. One can also speculate that if the Carnap of 1930 had been shown Lewis's more sophisticated analyses from four decades later, he would have immediately have accepted them as improvements. He might well have taken them as more sophisticated extensions of the analyses he was just beginning to produce himself.

It's also interesting to know how much Carnap's *Aufbau* influenced Lewis and his philosophical program. We know that Lewis read the *Aufbau*, because one of his first published work was



“Policing the *Aufbau*” in 1966, on an aspect of Goodman’s problem for Carnap’s phenomenal basis. However, he doesn’t much discuss other aspects of the *Aufbau* in his writing, and he doesn’t connect it to his central project of Humean supervenience. Still, it’s easy to surmise that it had a strong influence on the project.

Of course there are many differences between Carnap and Lewis. Perhaps the biggest difference is that Lewis is far more of a metaphysical realist. He is metaphysically inflationary, where Carnap is metaphysically deflationary. This applies to both the construction basis and the construction method. Lewis sees his basis as limning the ultimate metaphysical structure of the world, while Carnap sees it as the most helpful basis for science. Lewis sees the mereological ontology that drives his construction method as objectively true, while Carnap sees his set-theoretic ontology as pragmatically useful.

Still, it is interesting that one can abstract away from this metaontological difference between them and still see such commonalities between Lewis’ metaphysical construction project and Carnap’s pragmatic construction project. What for Lewis is the one true fundamental basis for reality is for Carnap an especially useful basis for many purposes, but many of the derivations may go similarly from there. Again and again, I find that both Carnap’s and Lewis’s views are powerful in ways that end up extending far beyond their own original commitments.

I would also like to ask: how would history have been different if Carnap had written his second *Aufbau*? Here I mean the work he might have written in the early 1930s, and not the Lewis-style extension. How would 20th century philosophy have been different if this *Aufbau* had been written? Certainly, the second *Aufbau* would have been a controversial work. It would have invited all kinds of disagreement. But it would not have been regarded as an obvious failure in the way that the first *Aufbau* has ended up being regarded. It would have been regarded as a serious viable work.

The first *Aufbau* was widely regarded as the last gasp of a dying research program, the program of phenomenalism. By contrast, I think the second *Aufbau* would have been the founding monument of a research program that was just being born, the program of 20th century physicalism. Of course, Carnap is widely regarded as the founder of 20th-century physicalism in any case. He and Neurath introduced the name, and his 1932 papers on the unity of science and on psychology in physical language were important founding works of 20th-century physicalism. Still, they are often ignored, perhaps because they are seen as focusing on limited topics in the philosophy of science and the philosophy of psychology respectively. I don’t think they are widely regarded as the major works of 20th-century physicalism.

I think that if Carnap had written the second *Aufbau* it would have been the central work of 20th-century physicalism. At least, it could have been. And it could have set the research agenda for today. It would still be agenda-setting in the way that Lewis's work is agenda-setting today. Again, it would have needed some development around the edges, the sort that Lewis and others eventually provided. But I think we would be talking about Carnap's second *Aufbau* as the central work in the program.

I also suspect that the course of Carnap's own reputation would have been very different if the *Aufbau* had been associated with the exploding program of physicalism rather than the dying program of phenomenalism and verificationism. Of course Carnap was always far more than a phenomenalist and a verificationist, but for a long time he was dismissed as one. With a successful physicalist *Aufbau*, he might even have been recognized as the leading naturalistic philosopher of the 20th century, a status more often associated with Quine.

It is also entirely possible that if Carnap had written the second *Aufbau*, the substantive course of philosophy in the 20th century would have been different. It's possible that the program of physicalist reduction associated with people like Lewis, Armstrong, and Jackson from the 1960s onward could have been greatly accelerated. Carnap was doing something very similar. With a concrete work like the second *Aufbau* on the table, work in this area might have been accelerated. Perhaps we would have seen versions of functionalism before the 1960s, for example. Of course all this is counterfactual speculation.

Finally, I'd like to know: why didn't Carnap write the second *Aufbau*? Maybe he was just bored with Aufbaus by this point. Maybe he was getting interested in other questions. Perhaps it was protocol sentences or logical syntax. Perhaps his interest in construction turned into an interest in the unity of science, resulting in his 1932 sketches of physicalist reduction rather than a full-blown construction project. I'd be very interested to know why he didn't write it, and whether he had any sense of what was lost.

## References

Bourget, D.J. and Chalmers, D.J. 2014. What do philosophers believe? *Philosophical Studies* 170: 465-500.

Carnap, R. 1928. *Der Logische Aufbau Der Welt*. Leipzig: Felix Meiner Verlag. Translated as *The Logical Structure of the World* (University of California Press, 1967).

Carnap, R. 1932. Die physikalische Sprache als Universalsprache der Wissenschaft. *Erkenntnis* 2:432-65. Translated as *The Unity of Science* (Kegan Paul, 1934).

Carnap, R. 1932. Psychologie in physikalischer Sprache. *Erkenntnis* 3:162-76. Translated as "Psychology in physical language" in (A.J. Ayer, ed.) *Logical Positivism* (Free Press, 1959).

Carnap, R. 1950. Empiricism, semantics, and ontology. *Revue Internationale de Philosophie* 4: 20-40.

Carnap, R. 1955. Meaning and synonymy in natural languages. *Philosophical Studies* 6(3):33-47. Reprinted as Appendix D of *Meaning and Necessity* (University of Chicago Press, 1956).

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

Coffa, A. 1991. *The Semantic Tradition from Kant to Carnap: To the Vienna Station*. Cambridge University Press.

Friedman, M. 1999. *Reconsidering Logical Positivism*. Cambridge University Press.

Goodman, N. 1951. *The Structure of Appearance*. Harvard University Press.

Howard, D. 1996. Relativity, Eindeutigkeit, and Monomorphism: Rudolf Carnap and the development of the categoricity concept in formal semantics. In (R. Giere and A. Richardson, eds.) *Origins of Logical Empiricism*, pp. 115-64. University of Minnesota Press.

Jackson, F. 1982. Epiphenomenal qualia. *Philosophical Quarterly* 32:127-36.

Lewis, D. 1972. Psychophysical and theoretical identifications. *Australasian Journal of Philosophy* 50:249-258.

Lewis, D. 1997. Naming the colors. *Australasian Journal of Philosophy* 75:325-42.

Lewis, D. 1971. How to define theoretical terms. *Journal of Philosophy* 67:427-46.

Lewis, D. Introduction to *Philosophical Papers*, volume 2.

Newman, M.H.A. 1928. Mr. Russell's causal theory of perception. *Mind* 37:137-48.

Pincock, C. 2005. A reserved reading of Carnap's *Aufbau*. *Pacific Philosophical Quarterly* 86:518-43.

Quine, W.V. 1951. Two dogmas of empiricism. *Philosophical Review* 60:20-43.

Richardson, A. 1998. *Carnap's Construction of the World: The Aufbau and the Emergence of Logical Empiricism*. Cambridge University Press.