Chalmers' Blueprint of the World

A critical notice of David J. Chalmers, *Constructing the World*. Oxford University Press, 2012. Pp. 480. ISBN 978-0199608577. £32.00/US$39.95 (hbk).

By Panu Raatikainen, University of Helsinki

1. Introduction

David Chalmers' new book, *Constructing the World*, is based on his 2010 John Locke Lectures in Oxford. The focus of the book is somewhat different from what many would expect from him. For those who are familiar with his joint paper with Jackson on conceptual analysis (Chalmers & Jackson 2001), on the other hand, the general approach of the book will have a familiar ring. According to Chalmers, it "would be a mistake to think of this book as intended mainly to provide a foundation for arguments about the metaphysics of consciousness. … It would be somewhat closer to the mark to think of this book as intended to provide a foundation for the ideas about two-dimensional semantics." (p. xxii)

The title of the book refers to Rudolf Carnap's 1928 classic *Der Logische Aufbau der Welt*, usually translated as either "The Logical Construction of the World" or "The Logical Structure of the World." To begin with, one thing should be clearly understood: Chalmers here is not defending any kind of idealism or social constructivism. As he puts it, "I am not really constructing the world, any more than Carnap was" (p. xvii). But as he goes on to clarify, "one can see the current book as trying to carry off a version of Carnap's project in the *Aufbau*: roughly, constructing a blueprint of the world, or at least

1

providing a vocabulary in which such a blueprint can be given. The aim is to specify the structure of the world in the form of certain basic truths from which all truths can be derived" (p. xvii-xviii). Carnap's project was to derive it all from a phenomenal basis, and in the end, a logical one. Chalmers thinks that this is impossible, and that one must expand Carnap's class of basic truths and change the derivation relation, adding, "But with these changes made, I think that the project is viable and that some of the spirit of the Aufbau remains intact" (p. xviii). Chalmers takes *Aufbau*'s specific phenomenalist project only as an example of what he calls "a scrutability thesis," and reflects on a number of possible such theses, defending some of them.

Chalmers begins the book with a reference, not to Carnap, but to Laplace and what is known as Laplace's demon: "Laplace suggests that given the right basic information, and sufficiently powerful reasoning, all truths about the universe can be determined. For Laplace, this basic information included truths about the fundamental laws of physics and truths about the location of all fundamental entities at a time" (p. xiii). The idealized subject of such complete knowledge is commonly called "Laplace's demon." For Chalmers, Laplace's thesis, too, is an instance of a scrutability thesis: "It says that the world is in a certain sense comprehensible, at least given a certain class of basic truths about the world. In particular, it says that all truths about the world are scrutable from some basic truths. This means roughly that there is a connection in the realm of knowledge between the basic truths and all the rest" (p. xiii).

In fact, however, Chalmers is not even arguing for the standard knowability thesis – the view that all truths are knowable – popular in semantic anti-realism and the epistemic theories of truth, for example (although he does say that his scrutability thesis is "at least a cousin," or "a plausible relative" of the knowability thesis (p. xvi)). Chalmers is much more liberal here, and simply stipulates, for example, an unlimited knowledge of physical truths across space and time for the idealized subject. So it is simply assumed that all

questions about the past, for instance, are "decidable."[1] Instead, Chalmers is defending a kind of reductionist program, although it is an epistemic rather than ontological reductionism (for the difference, see e.g. Silberstein 2002); hence the label "scrutability."

Indeed, he submits that the scrutability thesis can be viewed as a version of the thesis of the unity of science:

> "Overall, we can see scrutability as a weak sort of reduction, one that is compatible with various sorts of irreducibility that are manifest in science. … Correspondingly, the scrutability thesis can be seen as a weak sort of unity thesis, consistent with the various manifest respects in which science is disunified. It avoids the most prominent objections to classical unity theses, but at the same time it shares something of their spirit, and it can do at least some of the work that we might want a reductive unity thesis to do." (p. 309)

Chalmers further distances himself from Carnap's specific views, that is to say, from strong empiricism and verificationism: "I am not a logical empiricist or a logical positivist. I do not share Carnap's verificationism or his sometime phenomenalism. … My views on the mind-body problem would have horrified Carnap. … To oversimplify, one might say that where Carnap leans toward empiricism, I lean toward rationalism" (p. xviii). Still, Chalmers says that "[i]n many ways, Carnap is the hero of this book." But it is perhaps more a matter of taste, or at least of emphasis, how much the view that Chalmers defends really resembles the views of the historical Carnap.

Constructing the World is an ambitious book, and also a large chunk to swallow. It has nearly 500 pages, 8 chapters and even 17 excursus that explore all sorts of related issues.

---

[1] Nevertheless, Chalmers himself says that the scrutability thesis "can play some elements of the role that the knowability thesis has been used to play" (p. 26) – I must say, I fail to see how it could possibly serve, e.g., semantical anti-realism á la Dummett.

I am obviously not able to comment here on all of the book's themes, but must necessarily focus only on a few selected points and skip a number of others. For example, I have decided to be silent about the various issues related to the philosophy of language (e.g. sense and reference, wide and narrow content, rigidity, two-dimensional semantics, etc.) even if the book is loaded with material on this topic. Chalmers' views on these issues would deserve a separate treatment.

After a brief summary of some of the book's main ideas, I will concentrate on several critical points, which is obviously unfair – but there is simply not enough space to list everything that I find agreeable in the book.

2. Varieties of scrutability

A natural formulation of the scrutability idea that may first come to mind is that if the Laplacean demon knew all basic truths, it could come to know all other truths by inference from these basic ones. This is what Chalmers calls *Inferential Scrutability*. In this picture, though, the demon is part of the same world it is trying to know. And because of this, all sorts of paradoxes threaten (see p. xv). In order to avoid such problems, Chalmers suggests that we think of the demon as contemplating the universe conditionally, as follows. For any true proposition *p*, the demon would be in a position to know that *if* all the basic truths obtain, *then p*. Chalmers calls this formulation *Conditional Scrutability*. (p. xv)

Chalmers, however, wants to make one final change:

> "A key element of Laplace's idea is that the Laplacean truths are all the truths that the demon needs. No other empirical information is needed for the demon to do its job. Here, the idea is that to know the conditional above—if the basic truths hold, then *p* holds—the demon does not need any empirical information in the

4

background. In principle, the demon could know the conditional *a priori*: that is, the demon could know it with justification independent of experience."

This is *A Priori Scrutability*. There is a compact class of basic truths such that for all true propositions *p*, a Laplacean demon would be in a position to know *a priori* that if the truths in that class obtain, then *p* (p. xv).

Here then, we have Chalmers' favored derivability relation. But what about the scrutability base? We have already noted that like many other philosophers, Chalmers as well thinks that the scrutability thesis based on phenomenal truths is not plausible. Then again, a physicalist may submit that all truths are scrutable from a physical base. This is a natural suggestion, but Chalmers, for one, famous for his defense of the non-reducibility of qualia or phenomenal consciousness, and even property dualism, obviously cannot accept it. Chalmers submits, however, that, roughly, all truths are scrutable from physical *and* phenomenal truths. More exactly, Chalmers adds that the base still needs to be complemented, first with indexical truths (e.g. 'I am such-and such,' 'Now is such-and-such') and second with a totality or "that's-all" truth (which entails various negative truths such as 'there are no ghosts'). Chalmers thus argues that all ordinary macroscopic truths are *a priori* entailed by this class, which he abbreviates as **PQTI** (physical, qualia, that's-all, indexical). **PQTI** is the standard or default base for most of the book.

In the book's final parts, though, Chalmers considers narrowing the base. He also grants that his **PQTI** base for scrutability is not philosophically principled in the way that, for example, Carnap's and Russell's were. Chalmers even lists (pp. 22-23; cf. p. 380) five possible *principled* scrutability theses: the thesis of (1) Structural Scrutability (all truths are scrutable from structural truths); (2) Fundamental Scrutability (all truths are scrutable from metaphysically fundamental truths); (3) Acquaintance Scrutability (all truths are scrutable from truths about entities with which we are directly acquainted); (4) Primitive Scrutability (all truths are scrutable from truths involving only expressions for primitive concepts); and (5) Narrow Scrutability (all truths are scrutable from truths whose content

is determined by the internal state of the subject). Chalmers briefly argues for the last three of these, more tentatively for (2), and in a restricted form (see below) for (1). However, in what follows I will focus more on certain general strategies that Chalmers utilizes in arguing for these kinds of more specific scrutability theses rather than on the details of the various theses.

3. How radical is the scrutability thesis?

How controversial is Chalmers' scrutability thesis? If it is interpreted broadly enough, it is not at all radical. For example, if we set aside phenomenal consciousness (and it was in any case assumed that all truths concerning it are directly included in the base), it seems to be widely accepted that, even if not reducible (in the classical sense), all of the other facts supervene on the physical. But supervenience amounts to determination, and this means that there must exist – whether or not we now know them – some kind of "bridge laws" (perhaps only one-way conditionals (see below)). Chalmers agrees. If these bridge laws are not in principle unknowable (and there seems to be no good reason to think they are), an idealized subject could come to know them, and derivatively all higher-level truths, however non-reducible (in the traditional sense of "reduction").

For Chalmers himself, however, scrutability does not come so cheaply, because in his view – here he refers to Kim (and actually Kim's view is in turn apparently influenced by Lewis) – "[s]crutability does not allow bridging laws in the base" (p. 304). From Kim's perspective, bridge laws themselves are also among the targets of reduction and explanation, and cannot be uncritically assumed in reduction (see Kim 1998).

On the other hand, Chalmers does in fact sometimes refer to supervenience and the existence of bridge laws in defending the plausibility of the scrutability thesis (p. 278-9, 280). It is not clear to me how to relate these remarks to Chalmers' above-mentioned view. What is more, at one point, Chalmers makes a truly puzzling comment: "We can

also allow Q [phenomenal truths] to include *any psychophysical truths*: truths concerning lawful regularities and counterfactual dependence between the phenomenal truths above and microphysical or macrophysical truths" (p. 111). I wonder if I am missing something here: this seems to directly contradict Chalmers' view mentioned above. And the idea that such truths, which must involve both mental and physical concepts, would count as phenomenal (truths about *what it is like* to be a given entity), in any case sounds odd.

In what follows, I will assume Chalmers' "official" view to be the more demanding interpretation, meaning that the scrutability thesis does *not* allow bridge laws in the base: the view that everything, except perhaps phenomenal consciousness, supervenes on the physical is widely accepted in contemporary philosophy, and it naturally entails the existence of the relevant bridge laws; allowing bridge laws would thus make the scrutability thesis an easy consequence of this commonly accepted view. Defending it would hardly require a 500-page book. Hence, my interpretation has at least the virtue of keeping the thesis interesting.

4. Chalmers on Quine and analyticity

Chalmers announces that he will reply to the critical arguments of Quine: "Chapter 5 uses the framework to respond to Quine's arguments against analyticity and apriority" (p. xix); and again, "In chapter 5, however, I will suggest that an analysis in terms of scrutability provides the materials required to show where Quine's arguments go wrong" (p. 25-6). This promises a little too much, because in reality, of the battery of arguments that Quine has presented in different contexts against various explications of analyticity, Chalmers is only dealing with the arguments in the popular paper "Two dogmas of empiricism" (Quine 1951), and in fact, Chalmers focuses solely on the argument from holism presented in the final section of the paper. This is too easy a victory.

Chalmers immediately sets aside Quine's critique in the earlier sections of "Two dogmas," a critique based on the charge of circularity. But certainly such circularity worries are potentially also relevant to Chalmers' framework: if the claim is that all truths are *a priori* scrutable from a certain limited (e.g. physical, or observational) base, truths about meaning and analyticity also should be. If, however, all one can offer, as an analysis, would be other truths the concepts of which are intimately connected to the concepts of meaning and analyticity, but are no closer to the concepts of the base, the claim is problematic. But I digress.

The argument that Chalmers does consider is an early formulation of the argument based on confirmation holism ('the Duhem-Quine thesis'). Its principal target is the specific idea, popular among logical positivists, that the (cognitive) meaning of a sentence can be identified with its empirical content, meaning its observational consequences, as well as the attempts to explicate the notion of analyticity (and, consequently, the notion of *a priori*) with the help of this idea.

And the problem with this idea – as Quine famously emphasized – is that in science, isolated sentences of the more theoretical sort, and even whole theories in themselves, often entail no observational consequences without some auxiliary assumptions; moreover, different choices of auxiliary assumptions result in different consequences. Therefore not all apparently meaningful statements have determinate "empirical content" of their own – so the attempt to explicate the concepts of meaning and analyticity in terms of empirical content also fails.

Now, neither Chalmers' views of meaning nor his views of *a priori* have anything to do with this empiricist idea, so it is hardly surprising that his framework can avoid this line of Quinean criticism. Accordingly, dedicating a whole chapter to arguing that it can may appear a bit excessive.

5. Hard cases: mathematical truths

In chapter 6 Chalmers considers a number of hard cases for the scrutability theses: truths about mathematics, morality, ontology, intentionality, modality, etc. Many interesting issues arise here, but there is not enough space to discuss these numerous cases in any detail. I must, though, comment on the case of mathematical truths.

Chalmers begins with the assumption that all knowable mathematical truths are knowable *a priori* (p. 261). This is widely accepted, although not entirely uncontroversial, so let us continue. But what if mathematical truths exist which are unknowable? If there are such truths, they are not knowable *a priori* in particular. It is natural to think that Gödel's incompleteness theorems show that such truths exist. As Chalmers himself puts it, "Most obviously, certain Gödelian statements in arithmetic are not provable from standard axioms, and one might argue that they are not knowable *a priori*, even under an idealization."

Chalmers, however, thinks that the problem can be circumvented with the help of an infinitary inference rule, the "omega-rule", as did Carnap at one point (in *The Logical Syntax of Language* (1934/7)). The rule is the following: From premises $\phi(0)$, $\phi(1)$, $\phi(2)$, and so on, for every natural number, infer $(\forall x)\, \phi(x)$. Chalmers suggests that "one can idealize away from our capacity to consider only a finite number of cases at once" (p. 262). Note that this rule presupposes that the idealized reasoner can always know the truth of infinitely many of its premises. Chalmers writes:

> "Our contingent limitations stop us from following this rule, but we can conceive of a being that lacks these limitations. Consider a creature with an infinite capacity for parallel reasoning: when it has to evaluate the truth of $(\forall n)\, \phi(n)$, it simultaneously evaluates $\phi(n)$ for each n (perhaps thinking faster for larger n, to

9

ensure a bounded thinking time), and responds with the verdict "true" if and only if $\phi(n)$ is true for each n." (p. 262)

He then adds:

> "Of course this is a relatively taxing idealization, but it is not out of bounds. It is grounded in familiar cognitive capacities, such as the capacity to evaluate $\phi(n)$ for any given n, and is extended merely by dropping a limitation on parallel processing." (p. 263)

Well, yes and no! It is true that in the standard construction of the unprovable Gödel sentence, which is a true universal statement of the form $(\forall x)\, P(x)$, the property $P(x)$ is decidable, and we can always demonstrate $P(n)$, for any natural number n. However, such sentences are only the simplest among the infinitely many and increasingly complex arithmetical truths that are unprovable. And in more complex cases, where $P(x)$ may itself contain (unbounded) iterated quantifiers (e.g., $P(x) = (\exists y)(\forall z)\, R(x, y, z)$), we often cannot so much as prove $P(n)$, for all n, even if $(\forall x)\, P(x)$ is true. In other words, unlike with the simplest Gödel sentences, in general cases we cannot even approximate the truth, or consider it as decidable in the limit. That is, we do not on the whole have "the capacity to evaluate $\phi(n)$ for any given n."

Consequently, the introduction of the omega-rule amounts to trivializing the question of scrutability: it does not differ essentially from the assumption that a knowing subject can directly recognize all arithmetical truths. Moreover, there is, of course, more to mathematical truth than just first-order arithmetical truths. And for the theory of real numbers, not to mention set theory (even if we set aside famous undecidable problems such as the Continuum Hypothesis), nothing like the omega-rule is available. It would perhaps be wiser simply to leave mathematical truth outside all scrutability theses.

6. What is scrutable from what?

When Chalmers moves from the rather general PQTI base to his more specific views on scrutability, it is sometimes difficult to keep track of what is, in his mind, supposed to be scrutable from what. For example, on the one hand Chalmers writes:

> "*Microphysical expressions* ... are *not* plausibly conceptually fundamental, and can easily be approximately defined in other terms. Correspondingly, *microphysical truths* are plausibly scrutable from other truths … the truths in *P* (the class of microphysical truths) will be scrutable from a class of truths involving the *O*-terms in the Ramsifications of microphysical expressions. … It is arguable that through such a process, *P* [microphysical truths] will be scrutable from a giant Ramsey sentence specifying the network of microphysical entities and properties with *O*-terms including observational terms … along with *causal* and *nomic* terms and mathematical and logical terms." (p. 319-320; my emphasis)

On the other hand, he says shortly afterwards:

> "I think that it is plausible that at least in our world, all *causal* and dispositional truths are scrutable from *truths about laws* (laws that do not mention causation or dispositions, such as standard *microphysical laws*), and perhaps truths about chance, along with non-nomic truths." (7.6. p. 339; my emphasis)

And,

> "So I am inclined to think that at least where our world is concerned, the only nomic truths that need to be included in the scrutability base are statements of *fundamental laws*." (p. 339)

Does this not say both that *microphysical* truths are scrutable from certain observational, nomic and *causal* truths, and that all *causal* truths are scrutable from certain *microphysical* truths? This at least appears circular. To be sure, Chalmers allows for the

11

possibility of more than one unique scrutability base, "the existence of multiple scrutability bases," but this does not seem to be the explanation in this case, and one is left wondering whether everything is really in order here?

Further, are not ***fundamental laws*** also microphysical truths, and hence, according to Chalmers, scrutable from observational, causal and nomic truths? And since causal truths are supposedly scrutable from microphysical nomic truths, does this, all together, say – though expressed in a rather florid way – that all other truths are scrutable from observational truths alone (with indexical and that's-all truths)? Surely this cannot be what Chalmers really wants to claim. I must say I fail to understand the overall picture these passages, for their part, are intended to give.

7. Ramsification and Structuralism

Chalmers' scrutability framework is, in his own words (p. 362), "at least a close relative" of the famous research program in philosophy known as "the Canberra Plan,"[2] inspired by the work of David Lewis. Like the latter, Chalmers' approach makes heavy use of the technique of "Ramsey sentences," "the Ramsey-Carnap-Lewis method," or "Ramsification" (cf. Psillos 2006). Chalmers leans on this method explicitly, for example, both in arguing for the scrutability of microphysical truths with their theoretical terms from observational and causal/nomic truths (p. 320; see also above), and in defending versions of structuralism – and it is natural to interpret him as presupposing it in a number of places even if it is not always mentioned explicitly.

If T is a theory, and P_1, P_2 …, P_n are (some, or perhaps all – but typically the observational) predicates of T that are at stake, the Ramsey sentence T^R of T is obtained by first replacing all these predicates with second-order variables, and then, to the result

---

[2]   See Nolan 2010, Braddon-Mitchell & Nola 2009.

of this replacement, prefixing the existential quantifiers with respect to those second order variables. The original theory thus gets "Ramsified".

Ramsey sentences are often associated with different versions of structuralism. Because of the special status of phenomenal consciousness for Chalmers, structuralism cannot hold for him across the board. Still, outside this tricky area, Chalmers finds structuralism attractive: "My own view is that no robust version of structuralism along these lines is correct, because of problems associated with consciousness. … Still, to say this much is compatible with holding with Carnap that structural expressions play a special role in science and in communication" (p. 420). Chalmers also sympathizes with structural realism in the philosophy of science (see below): "At least some version of structural realism looks promising in light of the current discussion. …. The thesis that all truths are scrutable from structural truths and phenomenal truths is not implausible" (p. 422). Moreover, Chalmers' qualified structuralism plays a key role in his response to skepticism (15th Excursus).

Now Russell and Carnap advocated an extreme version of structuralism, and an extreme version of the Canberra plan with global Ramsification (all non-logical expressions get Ramsified) amounts more or less to the same thing. Such strong views are widely known to be problematic. Chalmers repeatedly mentions the "Newman problem," and agrees that it is a fatal obstacle for such views. However, he also states again and again that the Newman problem can be avoided by leaving some non-logical expressions un-Ramsified (p. 8, 21, 363, 409). I am afraid, though, that Chalmers greatly underestimates the scope and difficulty of the Newman problem. Moreover, he completely ignores another weighty problem for the Ramsey-Carnap-Lewis method, the objection which might be called "the Scheffler problem."

As to the Newman objection, we need to distinguish two uses, or two versions, of the objection. First is *the traditional Newman problem*, which was presented against the pure or logical structuralism of Russell (Newman 1928), and also applies against a similar

view entertained at some point by the early Carnap, as well as against the extreme version of the Canberra plan with global Ramsification. The problem is that the relevant specifications of the world are nearly vacuously satisfied, if only the world has a sufficient number of objects. This is the setting that Chalmers considers.

But second, there is the Newman objection as formulated in the more recent debate, *the contemporary Newman problem* (as one might call it). It was initiated by Friedman and Demopoulos (1985), and analyzed in detail by Ketland (2004, 2009) (see also Ainsworth 2009); the setting here is a bit more sophisticated, and the objection is not presented against global structuralism but against structural realism: in this formulation, not all predicates are Ramsified away, but only the theoretical ones.

Recall that structural realism is the view that we should not believe that our scientific theories can manage to discover the nature of the postulated theoretical entities, as standard scientific realism suggests; rather, we should only commit ourselves to the structural content of our theories. Yet according to structural realism, this amounts to more than only their empirical content, which is all that empiricist antirealism allows. Structural realism thus aims to present a plausible middle way between radical empiricism and full-blown scientific realism (see Ladyman 2008 for an excellent overview). The concept of the "structural content" of a scientific theory, essential for structural realism, obviously needs clarification. A standard way of doing this is in terms of Ramsey sentences; Worrall and Zahar (2001), for example, argue that the real content of a theory is exhausted by its Ramsey sentence.

Now the contemporary Newman problem is the following. Assume that a (consistent) theory is *observationally adequate*.[3] (Note that the extreme versions of structuralism do not require the latter: even the observational predicates get Ramsified. Here, on the other

---

[3]   How exactly "empirical adequacy" should be explicated here turns out to be more complicated than initially assumed; see (Ketland 2004, 2009).

hand, the observational expressions are left unchanged.) Then the corresponding Ramsey sentence is almost trivially true, provided that the domain only has a sufficient number of objects. Therefore, the objection continues, Ramsifying Structural Realism collapses into radical empiricism, and cannot supply a genuine middle ground between standard scientific realism and empiricist antirealism.

Chalmers – though he gives some relevant references in a footnote – in practice completely ignores the contemporary Newman problem. This is unfortunate, for it is this version that is most relevant to both the Canberra Plan, and Chalmers' own Scrutability Framework. One is left in total darkness with the question of how exactly one is supposed to avoid the problem.

Furthermore, as was noted above, there is another problem concerning the Ramsification approach, an objection that we may call *the Scheffler problem*. (It goes back to Scheffler 1963, 1968, and was revitalized in Raatikainen 2012.) Chalmers does not even mention it. The objection is as follows. Even if we do *not* Ramsify all, but only some (e.g. theoretical), predicates of a theory, the resulting Ramsey sentence is still in some cases *logically true*, although the original non-Ramsified theory is not. Therefore the assumption that the Ramsey sentence of a theory faithfully captures the real content or the cognitive content of the theory is not at all plausible.

These problems are serious indeed for the Ramsey-Carnap-Lewis approach, and also for Chalmers' scrutability framework, inasmuch as it essentially leans on it. They cast serious doubt over Chalmers' approach. Anyone who builds a whole philosophical system on Ramsification should deal with them in detail and argue convincingly that the particular system is not vulnerable to them. Chalmers instead seems to be confident that once he refrains from Ramsifying every predicate and Ramsifies only some, he is safe from them. But this is emphatically not the case. At least, without a much more careful explanation of how exactly Chalmers' scrutability framework could avoid these problems, it remains unclear whether the program can be carried through at all.

8. *A priori* entailments

Let us revisit *a priori* entailments and *a priori* scrutability, which play such an important role in Chalmers' construction system. Now, Carnap and Russell based their classic systems on explicit definitions: the concepts of the various areas of knowledge were to be defined in terms of concepts in the scrutability base. Chalmers, on the other hand, admits that such definitions are often neither available nor forthcoming, as many expressions simply do not seem to have a definition. Chalmers mentions "knowledge" as an example: as is well-known, its traditional alleged definition as "justified true belief" failed in the light of various counter-examples due to Gettier, and all attempts to formulate a new explicit definition have also arguably failed.

Nevertheless, Chalmers thinks that there can be *a priori* entailments and *a priori* scrutability which do not require definitions. He takes, again, "knowledge" and the famous Gettier cases as an example (cf. Chalmers & Jackson 2001, section 3): "In particular, it is striking that in many cases, specifying a situation in terms of expressions that do not include 'knowledge' or its cognates (synonyms or near-synonyms) enables us to determine whether or not the case involves knowledge" (p. 13). Chalmers generalizes that "Something like this happens throughout philosophy, psychology, and other areas. We are given a description *D* of a scenario without using a key term *E*, and we are asked to determine whether and how the expression *E* applies to it" (p. 13). Therefore he concludes that "*a priori* scrutability does not require definability" (p. 14).

Even if Chalmers' analysis of the case of "knowledge" here were adequate – for the record, I think it is not[4] – one may nevertheless doubt whether all the different relevant cases could work in the way that Chalmers sketches. In particular, the question of the *a priori* scrutability of the truths of some "higher-level" special science, for instance,

---

[4]   See Yli-Vakkuri, "Scrutability scrutinized" (forthcoming).

biology, psychology, or sociology, from physical truths, or, the *a priori* scrutability of some microphysical truths involving theoretical terms from observational and causal truths, looks problematic. For example, an average speaker, perfectly competent in everyday language and familiar with words such as "gold," but ignorant about physics and the fact, for example, that the atomic number of gold is 79, may be presented with various "scenarios" in terms of physics, and be completely at a loss as to whether "gold" applies in a given scenario or not. No amount of *a priori* conceptual analysis will get him or her out of this predicament.

At one point Chalmers switches the example from "knowledge" to "water", and refers to Jackson:

> "Jackson (1998) gives a nice model of the water/H_2O case, arguing that 'Water is H_2O' can be derived from microphysical facts using the *a priori* premise 'Water is what plays the water role' and the empirical premise 'H_2O plays the water role', which is itself derivable from microphysical facts." (p. 306, fn 4)

This is, of course, how the Canberra plan (and Kim) in general aims to achieve reduction ("functional reduction") without any bridge laws in the base. ("The water role", or whatever causal role, is then often explicated with the help of Ramsey sentences, with all their problems; but let us focus here only on the general intuitive idea.)

I am afraid I cannot share this optimism of Chalmers and Jackson. Even if we accepted that notions such as "the water role" and "the gold role" were sufficiently clear, determinate, and stable (I, for one, am not at all sure that they are), and that the relevant identities (e.g. "water is what plays the water role") were *a priori* known, the whole approach still seems to tacitly assume that the relevant causal role can be expressed in the vocabulary of microphysics. For how else could the empirical premise "H_2O plays the water role" be derivable from microphysical truths? But it is hardly uncontroversial that such causal roles can in general be expressed in the language of microphysics; apparently

17

they are often instead formulated essentially in the language of some (non-reducible) higher-level special science.

Moreover, "Water is H_2O" is a paradigm of a type-identity. There is, however, a wide consensus that in many inter-level cases, no such type-identities exist, and that some higher-level properties are multiply realizable. Chalmers grants this and mentions economics as an example (p. 302). He only adds that his scrutability thesis does not require definitions and bi-conditionals (equivalences), only weaker one-way conditionals (implications) from physical to economic (and other higher-level) truths (p. 302-3). As Chalmers concludes, "the scrutability of economic truths from microphysical truths is quite consistent with the multiple realizability of economic kinds."

However, this is disappointingly weak – being consistent is one thing, and being even distantly plausible is another. Merely noting the consistency is a far cry from an explanation of how this could be possible, not to mention from a convincing argument that this is indeed the case. If the scrutability base is not allowed to contain bridge laws, it is presumably not allowed to contain one-way conditional bridge laws from microphysical truths to higher-level special science truths in particular. But without something like them, it is very hard to see how the higher-level truths of a special science could be scrutable from microphysical truths. Mere gestures toward the case of "knowledge" and the Gettier scenarios are hardly sufficient.

To make the issue more vivid, let us consider, instead of economics, a more familiar setting, also dear to many Canberra planners, namely that of mental states such as beliefs and desires. (Let us again put off phenomenal consciousness, qualia, etc., which may require a different story.) Are truths about such mental states *a priori* scrutable from physical truths? According to the usual functionalist story, also popular among Canberra planners – sometimes Chalmers seems to agree; but whether or not he does is unimportant: this is just an example intended to illuminate the idea of functional reduction – these are (in a sense) "definable" in terms of perceptual input, behavioral

output, other mental states, and causal relations between these. (I am not myself sure whether beliefs and desires can actually be "defined" purely in terms of their causal roles, or whether their representational content must be taken as independently given – representational contents, though, seem to be just as much multiply realizable as causal roles, so this is not a relevant issue as regards the question of reducibility.)

Consider, for example, the following, admittedly simplified case. Assume Jane desires to please her boss, and believes the best way to do so now is to assent to his proposal. The desire and belief together cause Jane to assent to the proposal of the boss. (Let us assume that Jane's desire to please the boss is her only reason to assent to the proposal, and that this desire is not overridden by other desires.)

The corresponding counterfactual conditional

> "If Jane would not have desired to please her boss, she would not have assented to his proposal"

is then true, and hence, the desire can be taken to be a cause of Jane's action. Let us also assume that the desire is multiply realizable. Now contrary to the popular view, the subvening physical state which realizes the desire does not necessarily have the same causal role as the realized mental state. Assume the desire was actually realized by the physical property P_1, but that it could also be realized for example by the physical property P_2 (and many others). The relevant counterfactual conditional

> "If Jane would not have had P_1, she would not have assented to his proposal"

is then false: even if she did not have P_1, but happened to have for example P_2, she would still have assented to the proposal. The underlying physical state here is thus not a difference-making cause (cf. Raatikainen 2010, 2013; Menzies 2008, List & Menzies 2009). Further, the causal role of the realized mental state apparently cannot be *a priori*

deduced from the causal role of the underlying physical realizer P_1 (whatever it is). (Note, by the way, that not only the essential mental states, the belief and desire, are (presumably) multiply realizable, but here so is also the behavior output, Jane's assenting (cf. Jackson 1996).) Consequently it is not plausible that we could, by reflecting on various scenarios presented in terms of microphysics, tell *a priori* whether the relevant counterfactual claims about beliefs, desires and assent hold in them or not, and that we could infer the causal role of such mental states from microphysical truths.

In sum, the claim that truths of various special sciences would be scrutable from physical truths remains unsupported, and appears from this perspective a pure matter of faith. And this is a rather serious weakness in Chalmers' overall scrutability thesis.

Panu Raatikainen   *University of Helsinki*

Acknowledgements

References

Ainsworth, P.M. (2009). "Newman's Objection", *The British Journal for the Philosophy of Science* 60, 135-171.

Braddon-Mitchell, David and Robert Nola (eds.) (2009), *Conceptual Analysis and Philosophical Naturalism*, MIT Press, 2009,

Carnap, R. (1928) *Der logische Aufbau der Welt*, Berlin: Weltkreis. English translation: *The Logical Structure of the World*. Berkeley and Los Angeles: University of California Press, 1967.

Carnap (1937). *The Logical Syntax of Language*. London: Routledge.

Chalmers, D. & F. Jackson (2001). "*Conceptual Analysis and Reductive Explanation*", *Philosophical Review* 110: 315-61.

Demopoulos, W. and Friedman, M. (1985). "Critical notice: Bertrand Russell's *The Analysis of Matter*: Its Historical Context and Contemporary Interest". *Philosophy of Science* 52: 621–639.

Jackson, F. (1996). "Mental causation", Mind 105, 377-413.

Jackson (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Clarendon Press.

Ketland, J. (2004). "Empirical Adequacy and Ramsification". *The British Journal for the Philosophy of Science* 55: 409–424.

Ketland, J. 2009). "Empirical Adequacy and Ramsification II". In A Hieke & H Leitgeb (eds.) *Reduction, Abstraction, Analysis: Proceedings of the 31st International Ludwig Wittgenstein symposium in Kirchberg, 2008*. Lancaster: Gazelle books, p. 29-46.

Kim, Jaegwon (1999). "Making sense of emergence", *Philosophical Studies* 95, 3-36.

Ladyman, J. (2007). "Structural Realism", in Ed Zalta (ed.) *Stanford Encyclopaedia of Philosophy*. http://plato.stanford.edu/entries/structural-realism/

List, C. & Menzies, P. (2009). "Non-reductive physicalism and the limits of the exclusion principle", *Journal of Philosophy* CVI: 475–502.

Menzies, P. (2013). "Mental causation in a physical world". In Gibb et al. (eds.), *Mental Causation and Ontology*. OUP, 58–87.

Menzies, P. (2008). "Exclusion problem, the determination relation, and contrastive causation". In Hohwy & Kallestrup (eds.), *Being Reduced.* OUP, 196–217.

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

Nolan, D. (2010). "The Canberra Plan", in Graham Oppy and N. N. Trakakis (ed.), *Companion to Philosophy in Australia and New Zealand*, Monash University Press, Melbourne, pp. 98-100.

Quine, W.V. (1951). "The Two Dogmas of Empiricism," *The Philosophical Review* 60, 20-43.

Psillos, S. (2006). "Ramsey's Ramsey Sentences". In M.C. Galavotti ed., *Cambridge and Vienna: Frank P. Ramsey and the Vienna Circle*, pp. 67–90. Dordrecht: Springer.

Raatikainen, P. (2013). "Can the mental be causally efficacious?" In Talmont-Kaminski & Milkowski (eds.) *Regarding Mind, Naturally.* Cambridge Scholars Publ, 138–166.

Panu Raatikainen (2012). "Ramsification and inductive inference", *Synthese* 187, 569-577.

Raatikainen, P. (2010). "Causation, exclusion, and the special sciences", *Erkenntnis* 73, 349–363.

Scheffler, I. (1963). *The Anatomy of Inquiry*, Alfred A. Knopf, New York.

Scheffler, I. (1968). "Reflections on the Ramsey Method", *Journal of Philosophy* 65, 269-74.

Silberstein, M. (2002). "Reduction, Emergence, and Explanation". In P. Machamer & M. Silberstein (Eds.), *The Blackwell Guide to the Philosophy of Science*, Oxford: Blackwell, pp. 80-107.

Worrall, J. and E. Zahar (2001). "Ramseyfication and Structural Realism", Appendix IV, in Zahar, E.: *Poincaré's Philosophy: From Conventionalism to Phenomenology.* Chicago and La Salle: Open Court, 2001.