TWELFTH EXCURSUS

# Scrutability and the Unity of Science

The unity of science was one of the central concerns of the Vienna Circle. Otto Neurath edited the huge, never-completed *Encyclopedia of Unified Science* (see especially Neurath, Carnap, and Morris 1971). In his 1932 article 'The Physical Language as the Universal Language of Science' (translated into English as the 1934 book *The Unity of Science*), Carnap wrote:

> The opinion is generally accepted that the various sciences named [philosophy, formal sciences, natural sciences, social sciences] are fundamentally distinct in respect of subject matter, sources of knowledge, and technique. Opposed to this opinion is the thesis defended in this paper that science is a unity, that all empirical statements can be expressed in a single language, all states of affairs are of one kind and are known by the same method.

There is no single thesis of the unity of science. An *imperialist* unity thesis (embraced explicitly at the end of Carnap's article) holds that all sciences are part of a single science such as physics. A *reductive* unity thesis holds that all correct scientific theories in different domains are somehow reducible to or grounded in a single theory, such as a fundamental physical theory. A *connective* thesis holds that correct scientific theories in different domains have mutually supporting connections between them. A *similarity* thesis holds that correct scientific theories in different domains have some similarity in methods or form. A *consistency* thesis holds only that correct scientific theories should be consistent with each other.

The logical empiricists are often associated with imperialist or reductive versions of the unity of science thesis, although connective and similarity theses are also prominent in their writings.[1] In recent years, the trend among philosophers of science has been to reject strong unity of science theses in favor of weaker theses such as connective theses, or to argue that science is not unified at all. It

[1] For analyses of the logical empiricists on the unity of science, see John Symons et al., *Otto Neurath and the Unity of Science*. For recent work favoring the disunity of science, see John Dupré's *The Disorder of Things*, Nancy Cartwright's *The Dappled World*, and Galison and Stump's collection *The Disunity of Science*.

is widely accepted that attention to the scientific practice reveals far more diversity between the sciences than unity.

Scrutability has at least some bearing on the unity of science. The thesis that all truths are scrutable from base truths naturally suggests that all scientific truths are grounded in certain base truths. And the scrutability theses discussed here give a central role in the base to microphysical truths. So this might suggest a commitment to a strong, reductive version of the unity of science thesis. This could be read as a point in favor of the scrutability framework: it reveals a sense in which science is unified. Alternatively, it could be read as an objection to the framework: it is committed to a reductive thesis that the philosophy of science has revealed to be implausible.

The issues here are subtle, but it is worth exploring just what sort of unity thesis might follow from the scrutability thesis. For ease of discussion, I will start by assuming Microphysical Scrutability: the thesis that all truths are scrutable from the microphysical truths *P*. I will later consider how things are affected by a change to scrutability from *PQTI*. I will discuss both the constructive point (whether scrutability yields some form of unity) and the defensive point (whether objections to unity yield good objections to scrutability).

I will concentrate mainly on reductive unity theses. For any relation of reduction between theories, there is a corresponding unity thesis, holding that there is a single scientific theory to which all theories are reducible. But many notions of reduction and so of unity can be distinguished. One dimension of variation concerns which aspects of theories we are concerned with: their languages, their laws, their methods, their explanations, their true sentences? Another dimension concerns the character of the reduction relation: it might be logical (e.g., entailment), metaphysical (e.g., identity), epistemological (e.g., evidential grounding), or semantic (e.g., meaning equivalence). A further dimension concerns the structure of the reduction relation: it might be conditional (yielding one-way conditionals from the reducing theory to the reduced theory) or biconditional (yielding two-way conditionals from one theory to the other).

The Microphysical Scrutability thesis can be seen as a unity thesis involving epistemological conditional relations among truths: the truths in the languages of all correct theories are epistemologically deducible from the truths of fundamental physics. This thesis has a strong reductive flavor. But it does not entail the traditional unity theses that are now widely rejected.

One aspect of the classical conception of reduction is definitional reduction. Definitional reduction concerns semantic biconditional relations among language: the key claim is that the expressions of the reduced theory can be defined in terms of the expressions in the reducing theory. The corresponding classical unity thesis, found explicitly in Carnap's work on unity, is a definitional unity

thesis: the expressions of all correct theories are definable in terms of those of a single theory (such as physics).

The scrutability thesis does not entail the definitional unity thesis for a familiar reason: scrutability does not require definitions. A definitional reduction from economics to physics would require that economic expressions be definable using microphysical expressions, which in turn requires biconditionals connecting economics and physics. By contrast, scrutability requires only one-way conditionals from physical truths to economic truths.

This allows the scrutability thesis to escape perhaps the most well-known objection to classical unity theses: the objection from multiple realizability (e.g. Fodor 1974). On the face of it, economics could be realized in physics or in ectoplasm. Definitional reduction of economics to physics appears to rule out the possibility that economics is realized by anything other than physics. Furthermore, even in a physical world, different instances of an economic kind such as money might be grounded in a heterogeneous and open-ended class of physical realizations, suggesting that any physical definition would be wildly disjunctive. By contrast, the scrutability of economic truths from microphysical truths is quite consistent with the multiple realizability of economic kinds. In fact, the scrutability thesis can allow that in other scenarios, economic truths are scrutable from ectoplasmic truths.[2]

Another aspect of the classical conception of reduction is deductive-nomological reduction, often called Nagelian reduction after Ernest Nagel (1961). Nagelian reduction concerns logical conditional relations among laws. The key claim is that the laws of the reduced theory are entailed by the laws of the reducing theory, perhaps along with bridge laws. We might call the corresponding classical unity thesis a Nagelian unity thesis: the laws of all correct theories are entailed by the laws of a single theory such as physics, along with bridge laws.

The scrutability thesis does not entail the Nagelian unity thesis for a couple of reasons. First, the scrutability relation is weaker than logical entailment. Second, Microphysical Scrutability does not say that all truths are scrutable from microphysical *laws*: it says that they are scrutable from microphysical *truths*, including the distribution of microphysical items throughout space and time as well as microphysical laws. It follows that any true laws in chemistry, economics, and so on are scrutable from microphysical truths, but not that they are scrutable from microphysical laws.

[2] When generalized scrutability of B-truths from A-truths obtains, there will be at least approximate definitions of B-expressions using A-expressions. One might think that this is enough for multiple realizability to cause problems. In the case of economics and physics, however, we have scrutability but not generalized scrutability, precisely because there are scenarios in which economics is not grounded in physics. Scrutability alone does not support even approximate definitions of economic expressions in physical terms, at least if definitions are required to be a priori.

This allows the scrutability thesis to escape another objection to classical unity theses: the objection from contingency. On the face of it, there is contingency in biology or sociology that goes beyond the contingency of physics. The principles of neuroscience could easily have been different, even keeping physics fixed. Certain key constants of social network theory appear quite arbitrary. So these principles do not seem to be derivable from the laws of physics alone. To handle this problem, a Nagelian reductionist needs to allow initial conditions and not just laws in the reduction base. There is no analogous problem for the scrutability thesis, which has microphysical truths in the reduction base. The microphysical truths underlying brains and societies will themselves be contingent and arbitrary: even holding physical laws constant, they could have been different. And it is plausible that this contingency matches up well with the contingency of neurobiology and sociology. So for all this objection says, it remains plausible that neurobiological and sociological principles will be scrutable from all the microphysical truths in the vicinity of brains and societies.

Still, the scrutability thesis shares something of the spirit of the Nagelian unity thesis. A priori entailment has something of the spirit of logical entailment: both might be seen as a sort of deducibility. Microphysical truths go beyond microphysical laws, but only so far. If physics is deterministic, microphysical truths are themselves entailed by and scrutable from microphysical laws along with microphysical boundary conditions (the state of the universe at the Big Bang, perhaps). And even if physics is nondeterministic, microphysical truths will be scrutable from these things along with the values of probabilistic variables. So microphysical scrutability might be seen as sharing some of the attractions of this classical unity thesis, without some of its costs.

In one respect, the scrutability thesis is stronger than the Nagelian unity thesis. The classical thesis allows bridging laws in the entailment base: chemical truths are entailed by physical truths plus physical–chemical bridging laws. Scrutability does not allow bridging laws in the base: chemical truths are a priori entailed by microphysical truths. Where logical entailment is concerned, bridging laws play the helpful role of connecting vocabularies. Where a priori entailment is concerned, this role is not needed: truths in one vocabulary can be a priori entailed by truths in a quite different vocabulary. One might think of the framework as akin to one that requires the bridging laws to be a priori, except that as we saw in chapter 1, a priori entailment does not require explicit bridging laws or definitions at all.

This difference is a benefit rather than a cost of scrutability. As Jaegwon Kim (1999) has pointed out, allowing bridging laws makes the Nagelian conception of reduction much too weak. To see this, note that many mind–body dualists (including myself) allow that there are laws connecting physical properties to mental properties, so that mental truths will be logically entailed by physical truths plus psychophysical bridging laws. The Nagelian model appears to predict

that on this view, the mental is reducible to the physical. But such a claim obviously mischaracterizes the dualist view. The underlying trouble is that there can be laws connecting entirely distinct domains, each of which is irreducible to the other. So for a connection that deserves to count as reducibility, mere bridging laws do not suffice.

Scrutability invokes the much stronger requirement of a priori entailment, which brings with it a sort of epistemological deducibility of higher-level truths from lower-level truths. It is arguable that something like this is required to satisfy one key desideratum of reducibility: that any epistemologically brute facts in the higher-level domain be grounded in epistemologically brute facts in the lower-level domain. Allowing bridging laws subverts this desideratum. Bridging laws can themselves introduce brute facts, as the case of mind–body dualism suggests. By contrast, scrutability favors the desideratum, at least if we allow that a priori truths are never brute.

There are many different notions of reduction, and there is no point getting into a verbal dispute over what counts as 'reduction'. But the desideratum outlined above corresponds to at least one key notion of reduction, or one key constraint on such a notion. We might call it *transparent bottom-up explanation*: once one has spelled out the lower-level facts, the higher-level facts are rendered transparent. That is, there is no residual mystery about what the high-level facts are or about how the low-level facts give rise to them.[3] This sort of explanation is a goal of many reductive projects in science. A reductive project in chemistry can reasonably aim to ensure that once one has spelled out all the physical facts about an organism, the chemical facts are rendered transparent. If this project succeeds, we may not have explained why all the physical facts obtain, but given that they obtain there will be no residual mystery about why and how they give rise to the chemical facts.

Where scrutability fails, transparent bottom-up explanation fails. This is borne out by the mind–body case. Even after spelling out all the physical facts, the mental facts are not transparent, so there is a residual mystery about how the physical gives rise to the mental. The same applies to options that are intermediate between a priori scrutability and bridging laws. For example, one could appeal to a posteriori identities or a posteriori necessities connecting low-level and high-level domains. But even these leave an element of bruteness in an explanation. If one 'explains' consciousness by saying that it is identical to a certain neural state and leaves it at that, then one has not given a transparent bottom-up explanation. In effect, the identity claim plays the same sort of explanatory function as a bridging law in the case of mind–body dualism. When scrutability fails, there will be a priori coherent scenarios in which the low-level facts are as they are and the higher-level facts are different. These scenarios cannot

[3] This intuitive sense of 'transparent' should be distinguished from the technical sense in E14.

be ruled out by the low-level facts alone, so the low-level facts do not transparently explain the higher-level facts. Instead, one needs primitive interlevel bridging principles in one's explanatory theory.

Many cases of reduction involve interlevel identities: the reduction of water to $H_2O$ is one such. But in this case, the identity claim 'water is $H_2O$' is itself scrutable from lower-level truths.[4] In this sort of case, the high-level truths are scrutable from lower-level truths and are transparently explainable in terms of them. But when the identity claim is not scrutable in this way (as in the consciousness case), it effectively functions as a primitive claim in a bottom-up explanation, playing the same epistemological role as a brute bridging law. To remove this element of bruteness and achieve transparency, something stronger is required. Scrutability can naturally play that role.

It might be argued that scrutability is too weak for transparent bottom-up explanation, on the grounds that a priori entailment can connect distinct domains. For example, if mathematical truths are a priori, then they are priori scrutable from physical truths (or by any other class of truths), but they need not be reducible to physical truths in any reasonable sense. Likewise, two sets of truths can be a priori scrutable from each other, but it seems odd to hold that they can be reducible to each other. I think this is a reasonable criticism, and suggests that scrutability needs to be strengthened to yield the relevant sort of reduction. Here one might strengthen the requirement by moving from a priori entailment to the stronger sort of in-virtue-of claims discussed in chapter 1 and the sixteenth excursus (especially the conceptual grounding relation discussed there), or by moving to more specific models of scrutability-based explanations such as the mechanistic model that follows. In any case, scrutability will still plausibly be a necessary condition for a relevant sort of reduction.

Scrutability is also a weak constraint insofar as good reductive explanations require the low-level phenomena doing the explaining to have a certain internal unity. Scrutability could be satisfied even if microphysical truths were entirely chaotic, non-law-governed, and disunified; but in that case microphysical truths at best explain macrophysical truths in a weak sense. In the actual world, microphysical truths have a certain internal simplicity and autonomy that makes for better explanations than this, but the degree of simplicity will vary from case to case: a reductive explanation of the Second World War might be a poor one,

[4] Frank 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. This in effect invokes a definition of 'water' (although a functional rather than a microphysical definition) to ground the derivation. As always, the scrutability framework can dispense with the definition, but 'water is $H_2O$' will nevertheless be scrutable insofar as it is scrutable that $H_2O$ plays the key roles associated with water.

precisely because of the complexity of the microphysical base. Still, even when the base is arbitrarily complex, scrutability allows a sort of transparent bottom-up explanation: *given* the low-level truths, high-level truths fall out. Good reductive explanation requires something more, but scrutability will again be a necessary condition.

A model of reduction that is quite compatible with scrutability while imposing further constraints is one grounded in *mechanistic explanation*.[5] On this model, high-level phenomena are explained in terms of the orchestrated functioning of a mechanism: a structure performing a function in virtue of its components parts, component operations, and their organization (Bechtel and Abrahamsen 2005). For example, DNA and RNA molecules might serve as a mechanism by which the transmission of hereditary characteristics is enabled, thereby explaining genetic phenomena. Mechanistic explanation typically proceeds via functional analysis of high-level phenomena, casting high-level explananda in terms of functional roles. For example, the genetic phenomena that need to be explained are the functional roles of transmitting hereditary information. One then shows how lower-level mechanisms can play those roles and how DNA can transmit hereditary information. In this way, one achieves transparent bottom-up explanation.

Employing the scrutability model, we can divide this picture into three parts. First, high-level explananda are expressed using functional concepts, or concepts involving functional roles. For example, the concept of a gene can be seen as a concept of an entity that transmits hereditary characteristics in a certain way. Second, one tells a story about how low-level mechanisms play the relevant roles: about how DNA transmits hereditary characteristics, for example. Third, given that the roles in the mechanistic story and the functional analysis match up well enough, high-level truths will be scrutable from the mechanistic story. In effect, functional analysis grounds scrutability from underlying mechanisms.

I do not say that reductive or mechanistic explanation in science requires scientists to demonstrate an a priori entailment from low-level truths to high-level truths. That claim would be much too strong. Still, I think there is an important sort of reductive explanation in science for which scrutability is at least a tacit constraint. That is, it is a tacit desideratum that in principle, a given reductive story could be fleshed out with further lower-level truths, such that higher-level phenomena would be scrutable from there. If it turned out that such scrutability were impossible in principle, then the reductive explanation

[5] For my own version of a mechanistic picture of reductive explanation, see section 2 of 'Facing Up to the Problem of Consciousness'. The scrutability model is also compatible with other sorts of reductive explanation, including structural as well as functional explanation, but functional explanation by mechanisms is certainly the most common kind.

could reasonably be regarded as defective, or as failing to satisfy an important desideratum of transparency. In practice, reductive explanations typically proceed by giving just enough detail to make it plausible that a fleshed-out story of this sort could be obtained.

Some will think that scrutability is too strong a constraint on the grounds that the connections between physics and biology, say, are empirical rather than a priori. I have already answered this objection in arguing for scrutability. But it is worth keeping in mind again that scrutability does not require definition of biological notions in microphysical terms, and allows us to appeal to all microphysical truths and not just microphysical laws. And as before, even though bridging principles such as 'water is $H_2O$' are empirical, this is no bar to the a priori scrutability of the principles themselves from low-level truths.[6] For example, it remains plausible that someone using a Cosmoscope armed with all microphysical truths (along with phenomenal and indexical truths) could ascertain all the 'water' truths and all bridging principles connecting water and $H_2O$. There are tricky cases here, such as the interface between the quantum and classical domains, but these cases can be handled as in the discussion of macrophysical truths above.

Some may worry that other standard worries for Nagelian accounts of reduction will apply to scrutability. We have seen that standard problems tied to definability, to multiple realizability, and to bridge laws will not arise. Nor will problems tied to logic: some versions of a Nagelian account require that all theories be formulated in first-order logic, but scrutability does not. Another problem for Nagelian reduction concerns the 'reduction' of an old theory to a new one: the old theory contains falsehoods, which cannot be entailed by truths. The falsehoods in the old theory will not be scrutable from truths either, but various nearby truths will be, including claims that those falsehoods are approximately true, or true in certain circumstances.

Another worry concerns the autonomy of the high-level sciences. Cellular biology, cognitive psychology, economics, and paleontology are all enormously different from physics and from one another. It would be crazy to do cognitive psychology by doing physics. These fields have their own methods and their own conceptual and ontological frameworks. Perhaps most importantly, they all have a sort of explanatory autonomy: economic explanations are different in kind from microphysical explanations, and cannot begin to be replaced by microphysical explanations.

[6] Marras (2005) argues against models of reduction in terms of a priori entailment by arguing that bridge laws are empirical and known inductively. I hope it is clear by now that this argument involves a non sequitur.

Scrutability is quite consistent with explanatory autonomy. If an economic truth (say, about the financial crisis in 2008) is scrutable from physical truths, then a weak sort of explanation of the economic truths in terms of physical truths will be possible. Given that the physical truths are as they are, we will be able to derive the existence of the financial crisis and so reductively 'explain' it. But for most purposes this will be a much poorer explanation than an economic explanation (in terms of credit mechanisms, for example). The 'explanation' will presuppose an enormously complex set of physical truths. Even if these truths are grounded in laws and boundary conditions, the boundary conditions and perhaps the laws will involve much irrelevant complexity. This 'explanation' may have little predictive power and little practical use. By contrast, an economic explanation may be far simpler, more systematic, more predictive, and more useful.

In general, I favor explanatory pluralism: there are multiple explanations of most phenomena, and which explanation we choose depends on our purposes. There are causal explanations, historical explanations, reductive explanations, and many others. Reductive explanations are useful for some purposes, especially in trying to get a sense of how the world as a whole hangs together (how could there be economic phenomena in a physical world?). These explanations help to give us a unified picture of the world. But for most purposes, they cannot take the place of other explanations.

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. One might label it (as I do in *The Conscious Mind*) a sort of reductive explanation without reduction, where the relevant variety of reductive explanation involves transparent bottom-up explanation in terms of underlying truths. At least there is plausibly a notion of reductive explanation here, for which scrutability is a necessary condition.[7]

Correspondingly, the scrutability thesis can be seen as a weak sort of unity thesis that is consistent with the various ways in which science is disunified. It avoids the most prominent objections to classical unity theses, but at the same time 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.

Of course microphysical scrutability is false, at least on my view. I think that phenomenal truths, indexical truths, and a that's-all truth are not scrutable from microphysical truths. Correspondingly, I think that these are not explainable in terms of physical truths. But we can add these to the scrutability base, yielding

[7] For more on the relationship between a priori entailment and reductive explanation, see chapter 2 (sections 2 and 3) of *The Conscious Mind*, and section 6 of 'Conceptual Analysis and Reductive Explanation'.

the thesis that all truths are scrutable from *PQTI*. How does this alter the foregoing?

If phenomenal truths are not scrutable from the microphysical, this brings out a certain disunity of the sciences. If we equate scrutability with reductive explanation, then phenomenal truths will not be reductively explainable in terms of microphysical truths. Nor will truths whose scrutability requires phenomenal truths: perhaps mental truths, social truths, secondary-quality truths, and others. These truths will be explainable in terms of physics and phenomenology, but not in terms of physics alone. Something similar goes for truths whose scrutability requires indexicals. For example, objective physical truths may leave open whether water is $H_2O$ or XYZ, so that a fully transparent explanation of the truth that water is $H_2O$ requires an appeal to indexical claims about our location within the world. Likewise for the that's-all truth: positive truths can be reductively explained in terms of microphysical truths alone, but a full explanation of negative truths requires something more.

Despite this expansion, certain stripped-down analogs of many of these truths will be reductively explainable. As John Searle (1992) has noted, physics can explain the 'objective' aspects of heat and color, if not the 'subjective' aspects. Here we might think of a subjective truth (in the relevant sense) as one with a relevant dependence on phenomenal or indexical truths, and objective truths as those without such a dependence. One might define an objectivized notion of 'heat' solely in terms of an objective causal role (perhaps in terms of expanding metals and the like), leaving out any connection to experience. Then objective truths involving this notion might be reductively explainable in terms of microphysical truths. The same goes for an objectivized notion of 'water', which will apply equally to $H_2O$ and XYZ. Given that all truths are scrutable from *PQTI* and that scrutability entails reductive explainability, it follows that all positive objective truths will be explainable in terms of microphysical truths, and that all objective truths will be explainable in terms of microphysical truths and a that's-all clause. Here the objective truths might be seen as one version of Sellars' 'scientific image' (where subjective truths are part of the 'manifest image'). We will then have a strong unity of the scientific image so conceived.

Still, the scientific image so conceived may be a pale reflection of actual science. Consciousness, mentality, sociology, secondary qualities, and other subjective aspects of the manifest image are all subject matters for science. A unity thesis that covers all of these will need to have more than microphysics in the base: it will need phenomenal and indexical truths too. The role of indexical truths is relatively minor. The most important addition will be certain psychophysical bridging principles: laws, identities, or necessities linking physical properties to phenomenal properties. As long as phenomenal properties supervene

on physical properties, this addition will bring phenomenal truths into the fold.

We might put this picture by saying that all scientific truths are grounded in physics and psychophysics. The dual base here is less unified than a purely microphysical base, but it still allows a good deal of unification. We retain a unified scientific picture of the world grounded in a few fundamental properties linked by a few fundamental laws, albeit with more properties and laws than on the physicalist picture. This reinforces the moral of the prior discussion: while various strong unity theses fail, the scrutability framework supports at least a moderate and attenuated conception of the unity of science.