

Problems of Reduction

Jaegwon Kim

Philosophical Concept

Reduction is a procedure whereby a given domain of items (for example, objects, properties, concepts, laws, facts, theories, languages, and so on) is shown to be either absorbable into, or dispensable in favour of, another domain. When this happens, the one domain is said to be 'reduced' to the other. For example, it has been claimed that numbers can be reduced to sets (and hence number theory to set theory), that chemical properties like solubility in water or valence have been reduced to properties of molecules and atoms, and that laws of optics are reducible to principles of electromagnetic theory. When one speaks of 'reductionism', one has in mind a specific claim to the effect that a particular domain (for example, the mental) is reducible to another (for example, the biological, the computational). The expression is sometimes used to refer to a global thesis to the effect that all the special sciences, for example chemistry, biology, psychology, are reducible ultimately to fundamental physics. Such a view is also known as the doctrine of the 'unity of science'.

1. Reduction by derivation and definition

Unity and simplicity are often touted as important virtues achieved through reduction (see Unity of science). By reducing one domain to another, we show the reduced domain to be either part of the second or eliminable in favour of it. Depending on the nature of the entities reduced, reduction will, therefore, promote ontological or conceptual economy and unity. When numbers have been reduced to sets, numbers no longer need be countenanced over and above sets; numerical concepts can be explained in terms of notions involving only sets; and laws about numbers follow from principles about sets. There will often be an explanatory gain as well: when gas laws are reduced to principles of statistical mechanics, via the kinetic theory of gases, we have an explanation of why these gas laws hold to the extent that they do – why the pressure, temperature and volume of gases behave (roughly) in accordance with the gas laws.

Furthermore, reduction is sometimes thought of as a way of vindicating or grounding a possibly suspect domain of entities; if number theory is successfully reduced to logic, as Frege claimed, that would show number theory to be as firm and well-grounded as laws of logic. If the mental realm is reducible to the physical-biological domain, this would remove whatever problems and doubts might becloud the scientific status of the mental. Such are the thoughts and aspirations that inspire and sustain reductive projects and reductionisms.

There are apparently simple cases of reduction in which logical-mathematical derivation alone suffices for reduction. When Galileo's law of free fall or Kepler's laws of planetary motion are derived from the principles of Newtonian mechanics in conjunction with applicable force laws (see Mechanics, classical §2), they are reductively absorbed into a more general and comprehensive theoretical framework.

According to logical behaviourism, mental expressions are definable – that is, given synonymous translations – in terms of expressions referring to actual or possible behaviour ('behaviour dispositions'). Logical behaviourism, therefore, exemplifies an attempt to reduce the mental to the physical-behavioural through definition (see Behaviourism, analytic). Thus, one way to reduce a domain of expressions, or concepts, is to provide each expression with a definition couched solely in the expressions of the base domain. On the assumption that synonymous expressions refer to the same entities, a definitional reduction would also accomplish an ontological reduction: entities referred to by the first group of expressions have been shown to be among those referred to by the second group.

Reduction becomes more complex for domains that do not yield to direct derivational or definitional relationships, as in the case of reductions involving theories each with its own distinctive vocabulary, for example the reduction of thermodynamics to statistical mechanics.

2. The Nagel model of reduction

Ernest Nagel (1961) articulated a model of reduction for scientific theories that has served as the principal reference point in discussions of reduction (see Theories, scientific). The guiding idea of this model remains derivation, but Nagel saw that to derive laws of a theory from another which does not share the same vocabulary, certain 'bridge principles' must be assumed as additional premises. Consider two theories, T_1 and T_2 , each with its distinctive vocabulary, V_1 and V_2 , where T_2 is the candidate for reduction and T_1 the reduction base. Theories in this context are construed as sets of laws. Nagel's model requires that there be connecting laws, standardly called 'bridge laws', correlating terms of V_2 with terms of V_1 . What form do such laws take and how many are needed for reduction? The simple answer is that they must be available in sufficient numbers and be sufficiently powerful to enable the derivation of T_2 -laws from T_1 -laws. But that depends on the strengths of the specific theories involved. The only general requirement worth considering, therefore, is the following:

The condition of connectivity: For each primitive n -place predicate F of V_2 , there is a predicate G of V_1 , such that for any x_1, \dots, x_n (in the domain of entities covered by the two theories), it is a law that $F(x_1, \dots, x_n)$ if and only if $G(x_1, \dots, x_n)$.

Thus, each predicate of the theory to be reduced must be connected, via a *biconditional law*, with a nomologically coextensive predicate of the reducer. These bridge laws in effect licence the

rewriting of every T_2 -law as a T_1 -statement; in general, they enable the translation of T_2 -statements into the language of T_1 . Thus, the only difference between Nagel reduction and definitional reduction is that in the former the translation is underwritten by empirical laws whereas in the latter it is based on meaning equivalences.

There is a sense in which the condition of connectibility alone guarantees reduction: if any T_1 -rewrite of a T_2 -law is not derivable from the laws of T_1 as they stand, simply add it to T_1 as a new law (assuming it, as well as the original T_1 -laws, to be true). This would extend the base theory, but not its vocabulary; and the extension would be warranted since the base theory was incomplete in failing to capture a true law statable in its vocabulary: namely, the T_1 -rewrite of the T_2 -law.

It is evident that when a theory has been derivationally reduced, it is conserved as part of the base theory; its laws are shown to be 'derivative laws' of a more comprehensive theory. Moreover, the condition of connectibility ensures that the properties posited by the reduced theory find their nomic equivalents in the reduction base, with which they may ultimately be identified when a successful reduction becomes entrenched. Thus, Nagel reduction is a species of 'conservative' (or 'retentive') reduction, which conserves, and can legitimize, that which is reduced; it contrasts with 'eliminative' reduction which dispenses with what has been reduced.

3. Emendations to the Nagel model: the Kemeny–Oppenheim model

Nagel's model has been criticized on various grounds, of which the following three are perhaps the most important. First, biconditional bridge laws are too weak, and must be strengthened into identities if they are to yield genuine reductions (Sklar 1967; Causey 1977). We need, it is argued, the identity 'temperature = mean molecular kinetic energy', not just a law that merely affirms covariance of the two magnitudes. As long as the reduction falls short of identifying them, there would be temperatures as properties of physical systems 'over and above' their microstructural properties. Moreover such correlations ('nomological danglers') cry out for an explanation: why does temperature covary in just this way with mean molecular kinetic energy? By identifying them we provide a short and conclusive answer: because they are in fact one and the same.

Second, the reduced theory is often only approximately true, and its laws are derivable only under special simplifying assumptions (which, strictly speaking, are false). For example, in deriving the gas laws in kinetic theory of gases, one has to make various assumptions, such as that collisions between molecules are perfectly elastic, that molecules are point masses, and so on. And the laws hold only approximately and within a fairly narrow range of conditions. Third, it is often pointed out that the condition of connectibility, which requires biconditional bridge laws, is unrealistic and can seldom be satisfied. Thus, temperature cannot, it is claimed, be

uniformly correlated, or identified, with a single micro-based property; it may be mean kinetic energy of molecules for gases, but something else for solids or in vacuums.

As a response to the second criticism, it has been suggested (Schaffner 1967; Churchland 1986) that the reduction of T_2 to T_1 should require only that a *corrected version* of T_2 (or its 'image' in T_1), rather than T_2 itself, be derivable from T_1 . We may also explicitly allow special limiting assumptions of the sort mentioned earlier. But to what extent can we correct or revise a theory without turning it into another theory? Obviously, the required corrections must be such as to yield the necessary bridge laws or identities, but there is no guarantee that any reasonable amount of tinkering will suffice to accomplish this. This takes us back to the last of the objections mentioned above.

Kemeny and Oppenheim (1956) proposed a model of reduction that departs extensively from the Nagel model, giving us a broader conception of reduction that does not require connecting principles between the theories involved. The gist of the new model can be stated thus: T_2 is *Kemeny–Oppenheim reduced* to T_1 just in case all observational data explainable by T_2 are explainable by T_1 . This definition can, and should, be understood in a way that does not presuppose a context-independent theory–observation distinction or any special conception of explanation. The core idea here is that anything, be it an observation or a low-level empirical generalization or law, that the reduced theory can explain, or predict, should be explainable and predictable by the base theory. The base theory, then, is at least as powerful, as an explanatory and predictive instrument, as the theory being reduced, making the latter at best otiose. This means that the reduced theory is ripe for elimination – unless, that is, it is also Nagel-reducible, and hence conserved as a subtheory of the reducer. Notice that any case of Nagel reduction is also a case of Kemeny–Oppenheim reduction (at least, on the standard deductive-nomological account of explanation; see Explanation §2). Unlike the Nagel model, the Kemeny–Oppenheim model requires no direct relationship between the reduced theory and its reducer; in particular, it involves no bridge laws of any form. In fact, the reduced theory may be false and the reducer true. However, it assumes that the two theories concern the same domain of phenomena (at least, that the domain of the reduced theory is included in that of the reducer). In any case, the Kemeny–Oppenheim model allows the replacement, or elimination, of the reduced theory, and hence can serve as a model of 'eliminative' reduction.

4. Microreduction and microdeterminism

A pervasive trend in modern science has been to explain macrophenomena in terms of their microstructures, and reduce theories about the former to theories about the latter. Examples abound in the physical and biological sciences. The reduction of thermodynamics to statistical mechanics (see Thermodynamics), the reduction of optics to electromagnetic theory (see Optics §1), the successes in solid-state physics and molecular biology (see Genetics; Molecular

biology; Sociobiology), and a host of other examples appear to attest to the fruitfulness of microreduction as a research strategy (Oppenheim and Putnam 1958).

Both the Nagel and the Kemeny–Oppenheim models can be applied to microreduction. What is needed is the idea that one theory is *amicrotheory* in relation to another. The rough idea is that the microtheory deals with objects that are *proper parts* of the objects in the domain of the macro-theory. More specifically, the domain of the microtheory will include objects that are parts of the objects in the domain of the macrotheory; in addition it will include aggregates of these micro-objects, and aggregates of aggregates, and so on. And the objects of the macrotheory are identified with certain complex aggregates in this domain. Moreover, the microtheory has a set of properties and relations characterizing its basic micro-objects, and will generate complex properties for aggregates of these objects from the basic properties and relations. These micro-based properties (for example, mean molecular kinetic energy of a gas) of aggregative structures of micro-objects are important for reduction: in Nagelreductions, it is these micro-based properties that will be correlated (or identified) by bridge laws with the properties of the macrotheory; in Kemeny–Oppenheim reductions, they will be needed in providing explanations of the data explained by the macrotheory.

The metaphysical underpinning of microreduction is the principle of *microdeterminism* or *mereological supervenience*: properties of a whole are wholly determined by, or supervenient on, the properties and relations characterizing its proper parts. No matter how complex an object (say, a human being, a planetary system) may be, if you put together qualitatively indistinguishable parts in the same structural relationship, you will get an exact duplicate of that object (barring basic physical indeterminacies). It is the task of science, then, to identify a physically significant decomposition of a whole into parts and develop a microtheory about these parts that will explain why the whole has the properties it has.

5. Multiple (or variable) realization and reduction

One influential objection often deployed against the claim that a given theory is reducible to another is ‘the multiple realization argument’. This argument was initially developed as an objection against mind–body reduction, but has been used to argue for the impossibility of reduction almost everywhere (Fodor 1974). Consider a higher level property, such as pain: pain as a psychological state is ‘realized’ by diverse physical-biological processes in widely divergent organisms and structures. The neural realizer of pain in humans probably has little in common with its realizer in octopuses; nor can we a priori exclude the nomological possibility of pain in inorganic electro-mechanical systems. This means that there is no single physical-biological state which could be correlated, or identified, with pain in a biconditional bridge law, and this defeats Nagel reduction of any theory of pain to an underlying physical-biological theory, as well as reductive identification of pain with a physical-biological state. The same argument is often applied to biological concepts, such as ‘heart’ and ‘digestive system’, functionally defined

physical concepts like ‘carburettor’ and ‘thermometer’, and dispositional concepts like ‘transparency’ and ‘water solubility’.

Notice that the multiple realization argument does not touch Kemeny–Oppenheim reduction in general, but works only against Nagel reduction which, as we saw, aims to conserve the reduced higher level properties. In any case, one reply that is usually rejected is this: why not reduce pain to a *disjunction* of its diverse physical realizers? But given the kind of extreme diversity involved, such disjunctive states are unlikely to be projectible nomic properties and there cannot, it seems, be a unitary theory dealing with them (Kim 1992). Another, more plausible, way of dealing with the phenomenon of multiple realization is to lower our reductive aspirations: the argument perhaps shows that *uniform* or *global reduction* of pain is not feasible, but not that human pain, octopus pain and Martian pain cannot each be *locally reduced* to human physiology, octopus physiology and Martian electrochemistry, respectively (Kim 1992). Thus, multiple realization is consistent with (in fact, it arguably entails) the local reducibility of higher level sciences: for example, human psychology to human neurobiology, octopus psychology to octopus neurobiology, and so on. One possible difficulty with this approach is that pain, when considered multiply reduced to a set of diverse physical-biological bases, seems to lose its integrity as a single mental state; pain as such appears either eliminated or else remains unreduced outside the ontology of the lower level theories.

6. The primacy of physical theory

Physics is generally thought to be our basic science, the only science that aspires to ‘full coverage’ of all of the natural world. But what does this mean? Does it mean that all the special sciences – that is, laws and generalizations of these sciences and the properties posited by them – are reducible to the basic laws of physics and fundamental physical properties? It used to be thought that the primacy of physical theory was equivalent to global physical reductionism. This view will be rejected by most philosophers. Many now doubt the possibility of reduction in almost all areas of science, and downplay the scientific and philosophical significance of reductions (although of course no one has seriously suggested that we would ever actually achieve a single unified science in the vocabulary of basic microphysics). Some have argued that even where reductions are possible, we lose important explanatory information about phenomena in a given domain when we focus only on their microstructures and neglect the larger ‘patterns’ that emerge at macrolevels. These macropatterns are claimed to cut across microstructures, and be capturable only by higher level laws (Fodor 1974). These remain controversial issues, however.

Views of this kind resemble the position of emergentism (Morgan 1923) on higher level ‘emergent’ properties. Those who hold these or related views often try to explain the primacy of physical theory in a thesis of supervenience or determination, the claim that physical facts (including physical laws) determine all the facts, or that worlds that are indiscernible in respect

of all physical features are one and the same world (Hellman and Thompson 1975; Kim 1984). However, the relationship between reductionism and various forms of the supervenience thesis remains controversial and continues to be debated (Kim 1984).

References and further reading

- Benacerraf, P. (1965) 'What Numbers Could Not Be', *Philosophical Review* 74: 47–73.
(Raises important ontological issues concerning the set-theoretic reduction of numbers.)
- Carnap, R. (1938–55) 'Logical Foundations of the Unity of Science', in O. Neurath, R. Carnap and C. Morris (eds) *International Encyclopedia of Unified Science*, Chicago, IL: University of Chicago Press, vol. 1, 42–62.
(A classic positivist statement on reductionism.)
- Causey, R.L. (1977) *Unity of Science*, Dordrecht: Reidel.
(A useful and comprehensive account of reduction in science.)
- Churchland, P.S. (1986) *Neurophilosophy*, Cambridge, MA: MIT Press.
(Contains discussion of reductive strategies in cognitive neuroscience.)
- Fodor, J.A. (1974) 'Special Sciences – or the Disunity of Science as a Working Hypothesis', *Synthese* 28: 97–115.
(Presents the multiple realization argument against reduction and reductionism discussed in §5, and defends the autonomy of the special sciences.)
- Hellman, G. and Thompson, F. (1975) 'Physicalism: Ontology, Determination, and Reduction', *Journal of Philosophy* 72: 551–564.
(Presents a version of physicalism that rejects reductionism in favour of a supervenience thesis.)
- Hooker, C. (1981) 'Toward a General Theory of Reduction', *Dialogue* 20: 38–60, 201–235, 496–529.
(A detailed survey and critical examination of the literature on reduction up to the late 1970s.)
- Kemeny, J. and Oppenheim, P. (1956) 'On Reduction', *Philosophical Studies* 7: 6–18.
(Presents the classic replacement model of reduction as discussed in §3.)
- Kim, J. (1984) 'Concepts of Supervenience', *Philosophy and Phenomenological Research* 45: 153–176.
(Distinguishes various forms of supervenience and discusses the relationship between supervenience and reduction.)
- Kim, J. (1992) 'Multiple Realization and the Metaphysics of Reduction', *Philosophy and Phenomenological Research* 52: 1–26.
(An assessment of the multiple realization argument against reduction.)
- Morgan, C.L. (1923) *Emergent Evolution*, London: Williams & Norgate.
(A classic statement of emergentism.)
- Nagel, E. (1961) *The Structure of Science*, New York: Harcourt Brace.
(The classic source of Nagel's derivational model of theory reduction as discussed in §2, and which also includes discussion of emergentism.)

Oppenheim, P. and Putnam, H. (1958) 'Unity of Science as a Working Hypothesis', Minnesota Studies in the Philosophy of Science, vol. 2.

(A useful presentation of the doctrine of physical reductionism through microreduction.)

Quine, W.V. (1964) 'Ontological Reduction and the World of Numbers', Journal of Philosophy 61: 209–216.

(Reviews some ontological issues in the reduction of number theory, but the discussion has wider implications.)

Schaffner, K.F. (1967) 'Approaches to Reduction', Philosophy of Science 34: 137–147.

(Stresses the point that the reduced theory must be appropriately corrected.)

Sklar, L. (1967) 'Types of Intertheoretic Reduction', British Journal for the Philosophy of Science 18: 109–124.

(Argues that 'bridge laws' should be construed as identities.)

Taylor & Francis
Not for distribution