

B. Reductionism and Supervenience

Neven Sesardic

The philosophical debate about reductionism has a curious recent history. Within less than 40 years (from the mid-1930s to the end of the 1960s), philosophers of science have in succession defended and, indeed, taken seriously the following claims on the issue: (a) that reductionism is *a priori* true, (b) that it is contingently true, (c) that it is contingently false, and (d) that it is *a priori* false. Of these, (a) is now completely abandoned, (b) is moribund, (c) is presently a dominant view, and (d) is an influential and controversial position (see D. Davidson, 1970), but largely restricted to the philosophy of psychology.

The *a priori* truth of reduction is the ambitious, purely linguistic reduction advocated by logical positivists. Science was to be unified by translating the content of all "higher" sciences (sociology, psychology, biology) into physical language. This project has failed. A given theory can be reduced to another one only if there is a close match between fundamental concepts of the two disciplines. This is Ernest Nagel's (1961) "condition of connectability." However, advocates of reduction came to recognize that the link between the concepts, if it obtains at all, has to be contingent, rather than analytic, and established through empirical discovery, rather than through "meaning postulates." The argument that reduction is contingently true, however, encounters a formidable difficulty in satisfying the connectability condition: There are strong general grounds for not expecting property-to-property correspondence across different sciences. Put differently, things that are similar from the point of view of Theory T_1 need not, and as a rule will not, be similar from the point of view of Theory T_2 . But if taxonomies of two theories are so disconnected, the talk of reduction does not seem to make much sense. Psychological "pains" may not correspond to a simple category of neurophysiological patterns. Why is this so?

The so-called argument from multiple realization has been used extensively to establish the contingent falsity of reduction in different contexts (biology vs. chemistry, psychology vs. neurophysiology, sociology vs. psychology, economics vs. physics, etc.). According to this argument, any property of the "higher" theory is likely to be instantiated in so many different ways at the level of the "basic" theory that the resulting disjunctive and gerrymandered predicates will cease to represent natural kinds. Natural kinds are understood as classifications of real properties in the

world. Hence, terms that refer to natural kinds can appear in scientific laws and function in explanations. The property of having a particular molecular structure is taken to represent a kind (e.g., gold) and explain other associated properties, whereas having the properties of being shiny, yellow, and malleable are not thought to designate a natural kind. (The words *higher* and *basic* are used here in a neutral sense: The first refers to the theory whose reduction is being considered, and the second to the would-be reducing theory.) For example, there may be useful relations to be discovered about optimal foraging behaviors in general (see chap. 1). Instantiations of these patterns may be found in insects, birds, humans and other mammals. Clearly, the proximate mechanisms controlling the various behaviors are diverse, including chemical, visual, and rational modalities that do not easily map onto a physical kind of classification.

Problems do not disappear once one decides to dismiss reductionism as a simplistic and, in the end, untenable, philosophical position. Antireductionism has serious metaphysical and epistemological difficulties of its own. First, although rejecting thoroughgoing reductionism, a number of authors seem to want to preserve the idea that higher level facts are dependent on the lower level facts. Even after renouncing the hope of reducing, say, psychology to neurophysiology, one might still wish to assert that, if relevant neurophysiological facts are fixed, psychological facts are thereby also determined. The same holds for the relation between biological facts and physicochemical facts, and between properties of social wholes and psychological facts. Some philosophers follow this view all the way to a form of physicalism by insisting that all sciences are ultimately anchored in physics; their view is epitomized in the slogan, "No difference without physical difference."

A weak form of reductionism is sought by appeals to the relation of asymmetric dependence or, technically, "supervenience." Roughly, a family of Properties A is said to supervene on another family of Properties B only if it is impossible for two things to differ with respect to Properties A without also differing with respect to Properties B. The task of conceptually sharpening this preliminary definition has proved to be exceptionally delicate. Analytical attempts to make precise the shared intuition of this dependence relation led to the proliferation of incompatible definitions of *supervenience*. Worse still, most of these definitions (or all, as some think) failed by being too strong and entailing reductionism, or by being too weak to convey the meaning of intended ontological dependence (for an overview article, see Kim, 1990a).

Despite the success or failure to defend asymmetrical dependence, if so desired, the antireductionist would still face two other major difficulties; one having to do with explanation, and the other having to do with causality.

1. Explanatory Autonomy?

First, there is the problem of "the division of explanatory labor." For the reductionist, this was a nonproblem because of the belief that all properties of a "higher" science were always just a subset of the properties of the more basic science; in fact, all explanatory work was done with the conceptual resources of the basic discipline. Many antireductionists are prone to oppose this view by a contrary claim of the complete explanatory autonomy of higher theories with respect to the lower level ones. Relying on the essential incongruity of different taxonomies, they tend to argue that a question raised within a given conceptual framework can always be relevantly and satisfactorily answered without leaving that same framework. This would mean not just that higher sciences have no reason to fear preemption by more fundamental disciplines, but, furthermore, that they could afford not to care at all about what happens at the more fundamental level. This isolationist overreaction to the threat of reductionism is not viable. Cases will exist when two levels of explanation are insulated from one another due to our temporary ignorance of how to integrate them, but such examples do not support the strong philosophical claim for explanatory autonomy. (For an argument to this effect, concerning the relation of psychology and neurophysiology, see Kalke, 1969; H. A. Simon, 1970.)

There is evidence, too, that taxonomies of different levels are not always mutually independent. An antireductionism that required total disconnection between levels would be too strong. Concepts of higher sciences have occasionally been modified and redefined to adapt them to more basic categories. A simple illustration is the way chemistry handed over to physics the sovereignty over its fundamental concept, "chemical element." Elements are defined by the number of protons in the nucleus; as a consequence, it happens, somewhat paradoxically, that isotopes of the same chemical element differ in their chemical properties. Building on sporadic examples of this kind, and above all on the hope that similar instances will be massively forthcoming in the near future, some philosophers are trying to breathe fresh life into reductionism. Particularly in the

area of psychology, these philosophers are prophesying the transformation beyond recognition of its conceptual apparatus with the advance of knowledge about underlying brain processes, and that "mature" psychology will eventually become indistinguishable from neurophysiology (see Churchland, 1986). Such a sweeping optimism concerning the approaching fusion of higher sciences with their *antidisciplines* (the term for lower sciences in E. O. Wilson, 1978) lacks support and, to a great extent, remains a mere article of faith. If the declaration of full explanatory autonomy of the higher sciences was seen as an overreaction to reductionism, the contrary firm belief that higher theories will be "cannibalized" by the more fundamental ones may in turn be regarded as an overreaction to the previous overreaction. In the present context, steering away from the extremes appears to be the only justifiable course.

2. The Specter of Causal Inertness

Taking up the problem of causality, what first attracts notice is that it pulls in the opposite direction from the explanation issue. Most regard higher sciences as indispensable for the task of giving explanations. Explanation being interest-relative, it sounds plausible to argue that, when the event to be explained is stated in terms of a given theory, any attempt to explain it by switching to a deeper-level vocabulary must end up by answering a different question, or by "changing the subject" (D. Davidson, 1970). The concept of causality is different. Causal relations are detached from any pragmatic component: They obtain independently of one's interests, descriptions, and beliefs. With causality, we cannot so easily resort to the strategy that worked nicely with explanation—by making the analogous "division of causal labor" among various sciences and having pluralistic peace.

In fact, arguments have recently been produced for the view that all causal labor is actually done at the most fundamental, physical level, and that, consequently, the possession of other (biological, psychological, etc.) properties can have no causal impact whatsoever. The skepticism about the causal efficacy of "higher" properties flows from their being construed as dispositions with multiple physical realizability. To take a biological example, relatively greater fitness in an environment is the cause of greater reproductive success. Suppose fitness under given circumstances is increased by either being darker or by being stronger than one's conspecifics. Arguably then, the enhanced reproductive success of

some organisms is not caused by their fitness (the property of being darker or stronger), but by their being darker or being stronger. The basic idea is that multiply realizable properties are disjunctive properties, and that disjunctive properties are inappropriate candidates for causal roles. A suggestion that a combination of causally efficacious properties could be causally efficacious would strike some philosophers as akin to the joke of Richard Feynman, who once said to his audience: "The thing that holds you in your seats is a combination of gravity and politeness" (Feynman, 1985, p. 4). There is no scientific theory (i.e., physical etiquette?) investigating the causal powers of such a property.

Another argument to the same effect starts from two plausible principles; (a) the materialistic rejection of substance dualism, and (b) the causal closedness of physics (the belief that any physical event is exhaustively causally determined by the physical properties of its causes). On this basis, it is then inferred that there simply no causal work remains to be done by other properties.

Alternative solutions to the problem of causal inertness have been proposed (see Jackson & Pettit, 1990; Kim, 1990b; Le Pore & Loewer, 1987; Sober, 1984), but at the present stage of discussion, these issues remain highly controversial. Reviving reductionism may be a dead option, but the antireductionists have serious philosophical problems on their hands as well.

3. Five Interfield Dilemmas: A Recapitulation

To structure this fairly condensed discussion of reductionism and autonomy, Table 3.1 is offered as an overview of the different standpoints discussed here. The strongest antireductionist position is substance dualism (or perhaps substance pluralism). This position claims that the higher and the lower sciences are divided by an ontological barrier, and that they deal with completely different spheres of reality. It is in this way that mind was separated from brain in Cartesian dualism, life from nonlife in the *élan vital* theories, social wholes from psychological facts in the radical versions of methodological holism, and so forth. Substance dualism is a historically important view, but plays no role in the contemporary debate. A weak version of substance monism (*token physicalism*, in the philosophical jargon) is widely accepted.

A subtler form of dualism (*property dualism*) is a more serious contender. For instance, even after conceding that all psychological events

are neurophysiological events, some still deny that psychological properties are (only a subset of) neurophysiological properties. The strength of this move is best seen by comparing it with its contrary—property monism. Property monism is, in fact, equivalent to reductionism (i.e., the view that, despite the lower and higher disciplines having different sets of predicates, they both denote only one kind of properties—those of the lower discipline), and this view is highly implausible for the already adduced reasons. Today, everyone is aware that it is non sequitur to argue directly for property monism from the strong case for substance monism; thus the recognized implausibility of substance dualism is no argument against property dualism. In short, whereas the truth of substance dualism does lead to property dualism, the truth of substance monism implies nothing about the ontology of properties.

TABLE 3.1
Five Interfield Dilemmas

<i>Substance</i>	<i>Property</i>	<i>Interproperty Relationship</i>	<i>Explanation</i>	<i>Causality</i>
Dualism	Dualism	Supervenience	Autonomy	Inertness
Monism	Monism	Independence	Nonautonomy	Efficacy

The further question about interproperty relationship makes sense only under the assumption that there are two families of properties (property dualism). The philosophical scene is dominated here by the search for the supervenience relation. Of course, there are skeptics who think that the appeal to supervenience is "obscurantist in the extreme" (Schiffer, 1987, p. 154). Still others regard it as an attempt to have one's cake (to be openly antireductionist) and to eat it too (to preserve, nevertheless, the pet reductionist claim about the ontological primacy of physics). But the sheer bulk of work done on this topic across different theoretical fields (biology, psychology, sociology, ethics) testifies that such a pessimistic attitude may be premature.

The explanatory dilemma autonomy vs. nonautonomy is the most central question of all five. It is "semiconnected" to the issue of property ontology. That is, property monism (reductionism) entails nonautonomy: The existence of only one kind of property, that of the lower discipline, implies that the whole explanatory work is done "from below," regardless

of whether we are aware of it. But the similar inference from property dualism (antireductionism) to explanatory autonomy does not go through. The fact that different conceptual resources cannot be amalgamated into one whole does not mean that they cannot be fruitfully combined. Needless to say, the tendency toward making higher disciplines explanatorily autonomous and insulated from the lower level approaches is not based solely on the invalid inferential move from property dualism. Among factors motivating the isolationist attitude are reasons like the inertia of disciplinary compartmentalization, the fear of one's own field of study being appropriated by "intruders" from other sciences, and perhaps also the pure uneasiness about having to come to terms with fundamentally different theoretical approaches. These are sociological excuses, not philosophical reasons. Isolationist explanatory pluralism could legitimately result from genuine difficulties involved in a project of integrating different explanatory strategies.

Finally, the issue of causality raises a puzzling two-pronged argument that the further away one's explanatory account is from the physical level, the more distant is one from the truly operative causal features of the situation. Social sciences would fare worst under this picture because their categories would be at several removes from the uniquely causally efficacious physical properties. In the 1950s and the 1960s, philosophers would hardly worry about this: From the neo-Wittgensteinian perspective, dominant at the time, in social science there was no place for causal explanation (see chap. 1). In contrast to the natural sciences, the task of explanation was regarded here as limited to making actions intelligible in the light of agent's reasons, and any use of causal idioms in the context of humanities was regularly dismissed as a crude methodological mistake. However, the situation has dramatically changed. The arguments for the noncausal stance crumbled under criticism, and the search for causal mechanisms has become a standard aim in social science research. Therefore, the alleged causal inertness of nonphysical properties remains an unresolved philosophical problem.