Reductionism and Antireductionism

Thomas Nagel

New York University Law School, 40 Washington Square South, New York, NY 10012 ,USA

Abstract. Reductionism is the idea that all of the complex and apparently disparate things *we* observe in the world can be explained in terms of universal principles governing their common ultimate constituents: that physics is the theory of everything. Antireductionism comes in two varieties: epistemological and ontological. Epistemological anti-reductionism holds that, given our finite mental capacities, we would not be able to grasp the ultimate physical explanation of many complex phenomena even if we knew the laws governing their ultimate constituents. Therefore we will always need special sciences like biology, which use more manageable descriptions. There may be controversy about which special sciences cannot be replaced by reduction, but that there will be some is uncontroversial. Ontological antireductionism holds, much more controversially, that certain higher-order phenomena cannot even in principle be fully explained by physics, but require additional principles that are not entailed by the laws governing the basic constituents. With respect to biology, the question is whether the existence and operation of highly complex functionally organized systems, and the appearance of self-replicating systems in the universe, can be accounted for in terms of particle physics alone, or whether they require independent principles of order.

1998 The Limits of Reductionism in Biology. *Wiley, Chichester (Novartis Foundation Symposium 213) p 3-14*

The reductionist idea, which was already a gleam in the eye of the Presocratic philosophers, and which has been such a creative driving force in the history of modern science, is this: All of the complex and varied and apparently disparate things and processes that we observe in the world can be explained in terms of universal principles that govern the common ultimate constituents out of which, in many different combinations, those diverse phenomena are really composed. The idea is that there exists, in principle, a theory of everything, in the form of a theory governing the one thing or few things of which everything else consists. Let me refer to this as the *ultimate* level of explanation. Reductionism has two aspects: constitutive and explanatory. The constitutive thesis is that everything is made of the same elements; the explanatory thesis is that everything that happens can be given an ultimate explanation in terms of the laws governing those elements.

There are two kinds of antireductionism: epistemological and ontological—having to do, respectively, with what we can know and what is really the case. [4]Epistemological antireductionism holds that even if *in reality* everything is explained by particle physics, we cannot, given our finite mental capacities, grasp the ultimate explanation of most complex phenomena, and would not be able to do so even if we knew the law or laws governing their ultimate constituents. We are therefore constrained to make do with rougher explanations couched in terms that our minds can accommodate.

With regard to many of the things that go on in the world, epistemological antireductionism is uncontroversial. If there were a detailed causal explanation at the subatomic level of the movement of a pianist's fingers in the course of playing a Beethoven sonata, we would not be able to hold it in our minds. The same is true of the ultimate explanation of a thunderstorm, or of a rise in the kangaroo population of Australia. Analysed into their ultimate constituents, these phenomena are too complicated for us to handle. Moreover, ultimate explanation of the precise physical event in each case would not tell us what we want to know. The understanding we want requires explanations that refer to features of those events that we can observe or grasp—explanations in terms of music, or meteorology, or population ecology. The idea of the ultimate explanation of such things must remain, at best, a pure idea, and the same can probably be said of most familiar phenomena.

Still, the pursuit of reduction is often successful and it goes in stages. Even if there are some things that we can be sure will never be epistemologically reducible to the ultimate level, there are levels in between, such as those of traditional chemistry or atomic physics, and development can occur over time in our capacity to explain complex phenomena in more basic terms. The 19th century chemistry summarized in the periodic table of the elements was explained eventually in terms of a physics of atoms composed of protons, neutrons and electrons, and the descent to more basic levels has continued, so that reductions are epistemologically possible now which were not in the past. Even so, explanations at higher levels often remain practically preferable and for many purposes indispensable.

Others directly involved in these sciences will be discussing the epistemological problem, but I want to concentrate on ontological antireductionism, since it presents special philosophical problems. It can take more than one form, but it is always a claim about what the world is like, not just about certain conditions or limits on our knowledge of the world. What does the natural world, including ourselves, *really* consist of—as opposed to how we with our limited minds and practical needs find it convenient to describe it?

It is easiest to explore reductionism by considering its denial, and ontological antireductionism can take two forms—constitutive and explanatory—corresponding to the two aspects of antireductionism already mentioned. I won't say much about constitutive antireductionism. Its most famous example is the [5]psychophysical dualism of Descartes—the position that not everything in the world is constituted of the same basic elements, those studied by physics, but that there are non-physical events or things as well—conscious mental events and conscious subjects. That remains an interesting issue, but it is distinct from the broader question of the limits of reductionism in biology, which is our topic, so I shall leave it aside. Nor shall I discuss the form of vitalism that holds that all living organisms, conscious or not, contain a non-material vital principle in addition to the universal chemical elements. I suppose no one believes that today.

My topic will be the ontological issue between reductionism and anti-reductionism at the level of explanation. The ontological antireductionism I want to discuss has to do with the physical world, including the biological organisms in it, conceived as complex physical systems. It is the position that some physical phenomena, even though they can be explained in terms of principles that fit their specific features, simply do not have an explanation at the ultimate level—that is, in terms of the universal laws governing their ultimate constituents.

For the purpose of examining this type of view, I shall simply assume that everything from stars to organisms is composed of the same ultimate common constituents and that those constituents behave in conformity with universal laws. It is clear that for practical purposes, much of our understanding of living things must stay close to the biological level. Even though new levels of explanation become available over time, they do not necessarily result in the elimination of the old. For example, I gather that explanations of heredity in terms of classical genetics, descended from the Mendelian theory, are not about to be simply replaced by explanations in the language of molecular biology (Kitcher 1984). But whatever the epistemological situation, there remains a question about the ultimate facts: Is the fullest explanation of all physical phenomena, including those of biology, the one that refers only to the most basic and universal physical level?

The question is sometimes put by asking whether there are *emergent* laws of nature—laws that emerge only at a certain level of complexity and that govern the behaviour of complex systems without being derivable from the laws that govern their simple constituents. One of the problems with this view is that whenever we discover apparently emergent laws, provided they are sufficiently precise, we have to consider whether they are emergent only relative to our present physical theory, or more absolutely. Resting content with emergence is contrary to the scientific impulse: we tend to try to postulate previously unknown properties of the constituents which account for the laws reductively after all. This may or may not turn out to be possible, but the tendency shows that the reductionist aim operates as a norm setting the ideal form of understanding toward which we hope we are heading.

[6] Nevertheless, since emergence is the main alternative to reduction, it merits exploration. Emergence might rely on a supposition of indeterminism at the level of particle physics. According to this possibility, indeterminism in basic physics leaves some things unexplained which are nevertheless explicable by principles that govern phenomena at higher levels of complexity, and perhaps govern the development of complexity as well: so while the existence and behaviour of such higher-order systems is consistent with the basic laws of physics that govern their ultimate constituents, they are not causally determined by those laws to the degree to which they are causally determined by higher-order laws. Relative to basic physics, in other words, such systems are purely possible but vanishingly unlikely, yet they are predictable by reference to higher-order principles. I won't say more about this, both because the supposition of indeterminacy is problematic (it arises only at the level of measurement, not of the basic equations) and because I believe the problem of emergence can be posed even about a physically deterministic world.

The question is, what would it mean for a higher-level phenomenon to be explicable *only* in higher-level terms? Suppose that there is, whether we can know it or not, an ultimate explanation of every particular thing that happens in the physical world—an explanation in terms of the laws governing its basic physical constituents that makes it as determinate as it can be made by any higher-level explanation. Suppose that this is true of everything from thunderstorms to the performance of piano sonatas or at least the actual physical events by which a particular performance came about. Still, it might be the case that this ultimate explanation of the individual event in all its detail left something unexplained: namely why *that type of higher-level event* occurred—why the pianist hit those notes rather than why that specific chemical event occurred. To understand why he hit the notes we need a cruder level of description of the antecedent circumstances that would produce that type of result, even if the physical details had been different, in both the causes and the effects.

Of course if there is an ultimate explanation of a specific physical event, and that event *is* an event of a higher-level type, then in a sense the ultimate explanation explains why an event of that type occurred. But it does not explain why *some event or other* of that type occurred, and that is often what we want to know. Explanation has to seek its own level (Garfinkel 1981). Events of certain physically non-basic types, such as a tree shedding its leaves or a person paying a phone bill, can occur in many different ways—that is, they can be instantiated by many different complex events that are distinguishable at the basic physical level. To explain why *some* event of that type occurred, one must find conditions and principles that fit the type, and not just its specific instantiation on that occasion.

The fact that we are interested in the explanation of a higher-level event does not guarantee that there is one: If a father wants to know why all three of his sons were [7]killed in the war, there may be no answer except a separate explanation of how each of them was killed. But sometimes we do find higher-order laws governing complex events—in economics, population ecology, the psychology of conditioning, or (I am told) thermodynamics—without being able to derive them from particle physics. By itself, this doesn't show whether the higher-order laws are irreducibly emergent, or only emergent relative to our present knowledge. We have reason to *look* for reductive explanations, level by level, for such apparently emergent principles, since it is contrary to the spirit of scientific inquiry to leave black boxes permanently unopened.

Still, the possibility of explanatory irreducibility is relevant to the question of the limits of reduction in biology. If there really are biological explanations that accomplish something that cannot even in principle be done at more basic levels, chemical or physical, that would be an important limit to reductionism, not just a limit on our knowledge but a fact about the world. The issue has been discussed by philosophers under the heading of the status of the so-called 'special sciences' (Fodor 1974). Do the truths of all the special sciences—sciences that deal only with a restricted subset of natural phenomena, such as biology, psychology, or economics (if it is a science)—derive from physics? This question is particularly acute with respect to the life sciences; there seems less reason to doubt the derivability in principle of geology or astronomy from physics.

How could the following two propositions both be true?

(1) Every event that happens in the world has a fundamental physical description and a fundamental physical explanation.

(2) Some facts about the world do not have a fundamental physical explanation, but do have a higher-level explanation.

The answer is that they could both be true if the higher-level explanations depended on principles governing the relations between general types of phenomena or properties that were not subject to correspondingly general characterization in ultimate physical terms, even though each instance of such a phenomenon had a distinct ultimate physical characterization. Perhaps not all naturally important kinds correspond to kinds definable in basic physics. If that were so, the laws operating at the higher level could not be derived from corresponding laws couched at the fundamental level, even though each event falling under the higher-level laws could be given a separate ultimate explanation.

Every science seeks interesting general principles that support counterfactuals (i.e. statements about what would have happened if things had been different), not just detailed descriptions of individual events.

Science tries to discover how things work in general, so that if we know why something happened, that tells us that something like it would have happened under other, relevantly similar [8]conditions, but that if conditions had been relevantly dissimilar, it would not have happened. So the dimensions of relevant similarity are crucial in the formulation of scientific hypotheses. Are there, then, some sciences for which the relevant kinds or similarity dimensions are irreducibly higher-level and not equivalent to similarities in the terms of a more basic science—or the most basic and universal of all?

The clearest examples of important properties that have no basic physical equivalent are not natural but conventional. The contingency of conventions of language and money make it self-evident that there is no purely physical equivalent of someone's asserting that the Earth is round, or of someone's rent having gone up 20% in the past year. Explanation in psychology and the social sciences is completely dependent on generalizations that refer to such physically irreducible properties. Since conventions depend on the contingent practices of human societies, however, that doesn't tell us much about the issue of reduction in the natural sciences. If there are important natural kinds that are not equivalent to basic physical kinds, it will not be because human conventions determine the ways in which they are physically instantiated. Conventions do not carve nature at the joints.

One concept specifically relevant to biology, however, is that of a *functional* property, and functional properties too have no basic physical equivalent. For a very simple but still non-natural illustration, consider the property of being in 'shift' position, as this applies to any keyboard writing device. It is a functional property because it is defined in terms of its role in relation to the conditional inputs and outputs of the device, rather than in terms of its intrinsic physical character, and it explains why when you strike a key, you produce a capital letter. In an old-fashioned manual typewriter, it is physically realized by the raising of the carriage by a key that acts as a lever, so that the upper half of each key hits the ribbon. In an IBM Selectric, it is physically realized by a 180° rotation of the ball. In my word processor, God knows how it is physically realized. Now the functional properties of an artefact are not candidates for ontological irreducibility, because even though they are not equivalent to natural physical kinds, neither are they natural kinds of any other sort: *they don't really explain how the world works,* because artefacts are constructed by human beings to work as they do on the basis of the physical properties of their constituents.

The interesting question is whether there are functional properties *in nature* that do not correspond to natural kinds of basic physics, but that really do enter into generalizations that physics can't explain. There certainly are functional properties in nature, such as that of being a wing or an eye or a gene or a muscle, and epistemologically they are essential to our understanding. But at this point we run into a reductionist response parallel to the point just made about artefacts: namely that the functional properties of biological systems are the products of evolution by natural selection among alternatives whose availability is [9]antecedently given by the physical properties of their constituents. That a neuron or an eyeball or a gene will work the way it does is therefore just as strictly derivable from pure physics as is the operation of a typewriter. In both cases, the functional properties may be indispensable to our practical understanding, but the real causality is physical, not functional. The theory of natural selection means that all reference to purpose or function in nature is misleading, because it has to be taken as shorthand for an entirely mechanistic historical explanation.

That is the conventional wisdom, at least among the vulgar. But is it true? It follows from the application of an undeniably attractive reductionist world picture—the idea that evolution by natural selection completely bridges the gap between particle physics and biology (Kauffman 1993).

This makes evolutionary biology radically reductive, and much harder to really absorb than is generally realized, because we are allowed to keep our old purposive explanations, with the vague qualification that they all have to be reinterpreted in terms of survival value. But are there more specific reasons to be convinced that the range of biological possibilities available for natural selection is simply a logical consequence of the most basic principles of physics?

There are two separate questions here. First, do the basic laws of physics explain the functional organization of organisms, i.e. the principles which specify in higher-level biological terms what a given type of system will do under different conditions, internal and external? Second, is the coming into existence of such organisms either entailed or rendered probable or anyway not vanishingly improbable by basic physics, given the existence of the physical cosmos? The existence of typewriters is something whose *possibility* is entailed by basic physics, but typewriters exist only because people have made them. If the answer to the first question is yes, so that the possibility of the existence of eyes and ovaries is entailed by physics, then the full ontological reducibility of biology to physics depends on the reducibility of the evolutionary process, without recourse to additional basic principles of higher-level organization. And it depends not just on there being a full physical account of the exact course of events which has actually occurred, but an account of why some such development of complex systems was likely or not unlikely to occur independent of the precise initial cosmic conditions, physically specified. This would have to include, of course, an account of why physics makes likely the coming into existence of the most important functional property of all, the property possessed by certain complex systems of being *self-replicating*.

To sum up: Epistemologically and practically, there are obvious limits to the level of reduction we can hope to achieve in biology, or that would be useful. Higher-order understanding will remain indispensable. As for the ontological question—what the world is really like—we just don't know how far reduction may go, but we can be sure that the search for ever more [10]reductive explanations will continue to be one of the most powerful motives of science.

References

Fodor J 1974 Special sciences, or the disunity of science as a working hypothesis. Synthese 28:77-115

Garfinkel A 1981 Forms of explanation. Yale University Press, London

Kauffman S 1993 The origins of order. Oxford University Press, New York

Kitcher P 1984 1953 and all that: a tale of two sciences. Philos Rev 93:335-373

DISCUSSION

Rose: I wanted to query some unexplained terms. Both you and Lewis Wolpert (in his introduction) used the term 'level'. What constitutes a level? Do you see this as an ontological or epistemological description of an aspect of the natural world? Secondly, you made a distinction between the terms 'explanation' and 'description'. I don't understand why a basic physical account of a phenomenon is an explanation, whereas a functional account, for instance of the role of a wing, or of money, is only a description. And the third issue that I don't understand is why one should automatically assume that if you can describe a phenomenon in physical terms this is in some way more 'fundamental' than describing it in terms of functions of a system.

Nagel: It is an open question whether levels are epistemological or ontological. Take the familiar distinction between chemistry and physics: this is an example of a successful reduction. In advance of the explanation of chemistry by physics it was possible to distinguish between those two levels and perhaps not to know the extent to which the first could be explained in terms of the second. Thus it is possible that a distinction between levels can be epistemological, and it is possible that it can be ontological: only the success of a reductive explanation will answer the question.

With regard to 'explanation', I simply use the word 'fundamental' to express the idea behind reductionism: an explanation of an event in terms of the principles governing its most basic constituents—those that it shares with everything else in the universe—can be called 'fundamental'. It may well be that an explanation of this kind, even if it exists, is inaccessible to us. Even if it is accessible, it may be useless for the purposes we need it for. An explanation has to be tailored to what we want to know about the event. Taking up the analogy of the piano player, if we want to know why he played those notes rather than why the muscles of his hands moved in a particular way, then we will want antecedent conditions which, if they had been different at that level, would have led to a different outcome at the level of [11]description at which the explanation is sought. Explanation is correlative with the kind of description of the thing to be explained.

Bray: In simple terms, understanding is a human attribute, not something that is intrinsic to the system. When I understand something, it is because I have a small number of mental objects which I'm able to manipulate internally according to well defined rules. This process provides a simple explanation for what I observe: in that sense it is a language.

I would like to raise the question of whether our intellects are expanding into the realm of computers. Is this going to enlarge our capacity for understanding, and thus change the levels at which we view reductionism?

Noble: The sense I get, whether correctly or not, from what Thomas Nagel has said, is to wonder seriously about the initial distinction between the epistemological and the ontological. If we identify a type and it is such that we are forced to say that because of the nature of that type we can't refer to there being fundamental, lower-level explanations of what is going on because that type can't be analysed or referred to at that level, aren't we saying also that there is something there at the higher level? That is to move over into ontology: it actually exists.

Nagel: I used the analogy of the typewriter. We can say it printed a capital letter because the shift key was depressed, but that's not really how the universe works. Typewriters in shift position are not real aspects of the causal order. Not everything that we can say about the world has the kind of reality that science is after. There is a real question in biology about the explanatory status of the functional level of explanation, which is comparable. Of course wings exist and eyes exist, but I was just trying to explain the opposition between the point of view which says that while we have to see the world in this way it really is all physics, and the view that what we see here is a part of the way nature is carved up.

Gray: I want to go back to something that Lewis Wolpert said in his introduction, which I think is an important distinction to keep in mind, namely the contrast between that which the laws at a lower level permit at a higher level, and that which the laws at a lower level impose or require at a higher level. This tension between physics and biology, or between the physical law explanation and the functional explanation, centres exactly around that distinction. If you consider the way natural selection works, it can only work if you have variation in a manner which we can describe quite effectively in terms of information. The analogies about libraries and the information contained in DNA sequences are familiar to all of us. You can only have an information system if you have degrees of freedom. If everything were determined fully by laws other than the selection process that is acting upon that information system, you couldn't have the variability in which the information consists. Coming down to the physicochemical basics, in so far as I understand them, you can have selection [12]operating upon the bases that make up DNA precisely because the physicochemical constraints permit several possibilities at each point. Natural selection can then work on those possibilities without violating the laws of chemistry and physics, but by adding something to those laws which is not itself explicable in terms of those laws alone, but is explicable in terms of the interaction between organisms and their environment. You can of course then say, when you look at any specific organism and its interaction with the environment, that everything is going on in a way which requires, and for which there can be given, a detailed physical explanation of each event in that interaction. But, none the less, in order to understand how functions work, it is necessary to bring in these further understandings from an information perspective, and this is not itself a physicochemical perspective.

Quinn: As a practising scientist it has been my experience that sometimes things work and other times they don't. But reductionism is always the thing you try first—it is like Pascal's wager: if you believe that the universe isn't explicable, you're dead before you start. If you believe that it is explicable and you try and explain it you might fail, but at least it's a try. A well structured testable wrong idea is better than a vague vaguely right idea. The extreme opposite of reductionism is complete description: Linnaeus versus Darwin.

Wolpert: Are you really saying, no reductionism, no science?

Quinn: No. What I'm saying is that anybody coming into a new field tries reductionism first: sometimes it doesn't work and you then try an alternative approach.

Holmes: Concerning the problem of explanation, in physics and biology there are apparently rather different expectations as to what constitutes an explanation. In principle, at least in physics, everything should be reducible to equations of motion or field equations, which are predictive. In biology we only have one theory, a historical theory which is in no grand sense predictive. Having said this, it is clear that global prediction of complex physical systems is in fact illusory so that physics and biology are actually reduced to doing the same thing: namely, looking at subsystems and doing essentially what every engineer does, solving small systems pragmatically using free parameters. Ultimately, therefore, explanation is something we use in a local sense quite effectively, but in these global senses actually lacks meaning.

Wolpert: Can you give an example of this?

Holmes: Let's say we can use the Schrodinger equation to get as far as a hydrogen atom, but as a many bodied problem we already go wrong, so we give up. In biology, we know only of Darwin's theory and we cannot predict with this. Again, we reduce to doing something else: to explain by defining subsystems which you can analyse with a small number of parameters, each chosen for their usability or measurability.

[13] *Gray:* There was a recent paper about the changing lengths of lizards' legs which was precisely a predictive application of Darwinian theory (Losos et al 1997).

Williams: My point is very simple really. If we take our thinking back to the beginning of the Universe and start from the big bang, we have a problem with the development of it shortly afterwards in that it did not expand just radially, but it gained angular momentum, which is the introduction of turbulence. Where does turbulence come from? Your reductionist principles have gone, because you have no idea where turbulence comes from. If you start from a point in the big bang, everything goes outwards on a radius; nothing goes round. In turbulence, things are going round but where did such motion come from? The point follows that if there is a one-off event, then we have a hard job conducting a scientific investigation or discussion of it. This means we can't reduce past the one-off events: the big bang, the introduction of turbulence, the origin of life.

Nurse: This issue of explanation versus description is interesting. I think of it differently to Thomas Nagel. For me there is often a greater sense of explanation when working at a higher level and more a sense of description when working at a lower level. Discussing a wing in terms of its function provides more of an explanation of a 'wing' than describing its molecular structure. This may reflect how humans think: that is, we only find satisfactory explanations when thinking about a phenomenon at levels close to the level at which the phenomenon operates.

My second point has to do with this issue of new laws coming into place at new levels. You mentioned the concepts of souls and of vitalism. There can be a lot of baggage that comes along with concepts such as these: when the word 'soul' is mentioned one immediately starts thinking of things that are not explicable by the laws of physics.

Morgan: My comments follow on because they are concerned with the cognitive issue. I wonder whether we should try to dissect this epistemological issue further. The proposal is that there are limits to reductionism that are imposed by the nature of our cognition and our minds. However, these could be of two very different varieties. One could be just to do with the amount of information that we can process: taking your example of a thunderstorm, suppose in principle that we could describe a thunderstorm by the movement of every particle within it. Clearly one reason why that would be useless is that there would be too much information for us to take on board. On the other hand, suppose one could derive a satisfactory equation describing those particles, that might still defeat our imagination in the way that natural selection defeated the imagination of many of Darwin's contemporaries, and the way in which the general theory of relativity still defeats the imagination of most people. These are entirely different forms of cognitive limit. Which of the two did you have in mind when you proposed epistemological antireductionism?

[14] *Nagel:* The epistemological/ontological distinction has been at the centre of much of this discussion. It is true that we don't want an understanding of everything in terms of particle physics. We wouldn't feel we understood biological processes if someone were to give us a particle physics account of them because we don't see the world in those terms: instead, we want to understand the world in the categories in which we perceive it and carve it up naturally. But I still think there remains the question of whether the real truth about causal order of the universe is somewhat different from what is easiest for us to grasp. Thus there might a reductionism which simply doesn't conform to the shape of the science that makes things comprehensible to us.

Reference

Losos JB, Warheit KI, Schoener TW 1997 Adaptive differentiation following experimental island colonization in *Aniolis* lizards. Nature 387:70-73