On the Way to a Theory of Antireductionist Arguments

PAUL HOYNINGEN-HUENE

The title of this paper indicates that it deals with work in progress. First of all, I will have to explain what I mean by a theory of antireductionist arguments. Having done this I will present some preliminary results concerning this theory. In the initial section, I shall try to awaken the need for a theory of antireductionist arguments. In the second section, I will discuss the type of theory which could be meant, and how we could obtain such a theory. This will explain why case studies of a certain kind are needed. In the third section, I will very briefly sketch such a case study. Finally, I will present some preliminary results concerning a theory of antireductionist arguments.

1. The Need of a Theory of Antireductionist Arguments

Let us first focus on the very idea of a theory of antireductionist arguments. In order to awaken the need of such a theory, I have to enter into the larger context of reductionism. It is well-known that in the history of science and philosophy there have been many heated and sometimes extremely long disputes about reductionism. Already Aristotle had to defend his cosmology, that postulated a fundamental difference between the celestial and the terrestrial region, against the opposite, reductionistic idea. In modern times, there have been many similar debates, such as those concerning the reduction of the mental to the physical, of thought to the perceived, of the electrical and the thermal to the mechanical, of the living to the non-living, of the chemical to the physical, of the mathematical to the logical, of the logical to the psychological, etc.

Some of these debates have apparently been settled for good, for instance the question of whether electromagnetic phenomena are mechanical in nature; there is a firm consensus in physics that this is not the case. But the majority of the disputes mentioned have turned out to be amazingly
long-lived. Examples include the question of the reduction of the living to the non-living, and the problems concerning the relation of the mental to the physical which is the main topic of this volume. There are several reasons for the longevity of controversies about reductionism. I will now sketch a few of them, restricting myself to the cognitive realm, and omitting sociological or psychological reasons which undoubtedly also play a role, at least sometimes.

For starters, the fact that the concept of reduction is extraordinarily ambiguous is a weighty reason for much confusion in the debates about it. Every discussion about reduction has to clarify its subject matter in the very beginning by drawing some distinctions. Of course, I cannot here address all the details and the still remaining problems of these distinctions, which have been extensively, although perhaps still not exhaustively discussed in the literature (see e.g. Ayala 1974; Ayala 1989; Fodor 1974; Hoyningen-Huene 1985; Hoyningen-Huene 1989; Nagel 1961; Nickles 1973; Oppenheim/Putnam 1958; Schaffner 1967; Schaffner 1976; Sklar 1967; Wimsatt 1979).

A first fundamental distinction concerns the question of whether or not one deals with reductions concerning essentially the same domain. For instance, reductions of successive theories deal with essentially the same domain (pace incommensurability: see e.g. Hoyningen-Huene 1990); the reduction of special relativity to classical mechanics is an example. Different domains are involved when we are dealing with reductions within a certain hierarchy of levels, for instance between physics, chemistry, biology, psychology, and sociology (there is, by the way, no consensus in the literature which the relevant levels exactly are). But the different domains may also be spatially separated. The question of the reduction between the celestial and the terrestrial domains provides a case in point which, for quite some time, was a controversial problem.

In the following, I will focus on reductions concerning different domains. In these situations in which a domain B is to be reduced to another, 'lower' domain A, at least four different concepts of reduction can be distinguished, roughly in the following way.

Ontological reduction concerns the question of whether the domains A and B are ontologically identical, i.e., whether B is constituted by the same elementary entities with the same elementary interactions as A.

Epistemological reduction concerns two questions: whether the concepts necessary for the description of B can be redefined (in an extensionally equivalent way) in terms of the concepts of A, and whether the laws governing B can be derived from those of A, supplemented by the redefinitions mentioned, suitable initial and boundary conditions, and possibly by suitable supplementary assumptions.
Explanatory reduction concerns the question of whether for every event or process in B there is a mechanism belonging to A which explains the event or process.

Methodological reduction is, in my opinion, best construed as a residual category; it concerns all questions that are not covered by the previous categories. For instance, heuristic research strategies may or may not be reductionistic. Further distinctions can be drawn, which are without interest in our context, however.

In the following, I will exclusively concentrate on epistemological reduction. However, a more than cursory discussion of epistemological reduction is today in need of justification, since some philosophers have argued that questions of epistemological reduction are uninteresting, and even quite misleading with respect to the actual practice of science (for example Hull 1973, p. 622; Hull 1974, pp. 12, 44; Maull 1977; Schaffner 1974; Wimsatt 1979; Wimsatt 1980). Now it seems to me that an exclusive occupation with questions of epistemological reduction so characteristic for the heyday of neopositivism, can indeed produce a very distorted image of science. The variety of problems of reduction in the actual sciences simply cannot be reduced to the single question of epistemological reduction. But the opposite extreme, namely the radical banishment of all questions concerning epistemological reduction, seems to be also mistaken, for epistemological reduction deals with the question of the conceptual distance between different branches of science. Even if this question were without any scientific import (which it is not), it would be still interesting from a philosophical point of view.

What can we say in principle about epistemological reduction? Clearly, there are the positions of reductionism and anti-reductionism which assert or deny the possibility of the respective reduction. But sometimes, there is the possibility of mediating positions, namely, if from a certain perspective, A and B appear as different aspects of the same subject matter. For instance, the question of the relation between the terrestrial and the celestial domain has, in modern times, not been answered by a reduction in one or the other direction, nor with a demonstration of the impossibility of a reduction. Rather, classical physics provides a point of view from which the opposition between the two domains dissolves, both with respect to the substantive and to the conceptual divergences.

Speaking as abstractly as before, the mere articulation of reductionist, antireductionist, and mediating positions does not seem to raise any problems of principle. But, in fact, there is a host of problems concerning the precise articulation of these positions. As soon as one asks questions about the set of laws involved, about the possibility of their logical articulation, about the sort of inferences admitted, about the premises
allowed in the derivations, about the nature of the sought redefinitions, and other related questions, the simple opposition among the principal positions disappears. Still, these positions seem to remain useful as guidelines in the sense of ideal types. At any rate, the possible transitions between these principal positions have further contributed to the confusion in the debate about reduction.

Now, if reductionist, antireductionist, and mediating positions can be distinguished, at least as ideal types, then the question arises how we can argue in support of these positions. What is the structure of an argument supporting either reductionism, or antireductionism, or the mediating position? This question can be clearly answered, at least in principle, with respect to the reductionist and the mediating position, respectively. These positions contain essentially existence statements: the claim that a certain reduction or a certain mediating position does indeed exist. In the empirical sciences, such existence claims are usually argued for constructively: by constructing the entity whose existence is claimed. (In the formal sciences, there are, under certain conditions, also other strategies for the existence proofs but this is without relevance here.) In other words, the reduction or the mediating position must be presented, with all the accomplishments they claim. In practice, however, such constructions can be seldom carried out in every detail, and one has to remain satisfied with a certain plausibility of the possibility of the construction, instead of the construction itself. But in principle, the structure of the needed arguments is clear.

The situation is quite different with respect to the antireductionist position. Here we seek not an argument for an existence claim, but for an impossibility claim: the reduction of B to A is impossible, or: all attempts to reduce B to A are bound to fail. It seems, at least in the empirical sciences, that an argument for such a universal statement can never be cogent nor is its structure clear. This is another reason for the longevity of debates about reduction: It is not at all clear what a successful argument for an antireductionist position looks like. Correspondingly, a good deal of the literature on reduction, both on the philosophical and on the scientific side, is somewhat unsatisfactory. Arguments may seem somehow plausible but they are not fully convincing; they may seem absurd but for unknown reasons; they may seem shallow but still possess great factual impact. In this situation, one may wish better to understand the antireductionist arguments themselves: not to attack or defend this or that antireductionist position, but to reflect on the arguments to be used or to be avoided thereby. In other words: the desideratum, then, is a theory of antireductionist arguments, with a sense of 'theory' yet to be clarified.

It is surprising that there is as yet no attempt in the literature to obtain a theory of antireductionist arguments, to the best of my knowledge,
despite the strong interest in the results of antireductionist arguments. In
structure, this situation seems similar to the situation of argumentation in
general before Aristotle. There were arguments which were more or less
successful, but the foundations of argumentation were obscure. It was,
then, an eminently philosophical task to reflect about the arguments
themselves, i.e. to invent formal logic. Of course, the problem of argu-
ments in general is immensely more important than the problem of
antireductionist arguments, but the structure of the problem is the same.

2. The Theory of Antireductionist Arguments

Se let us move towards the envisaged theory of antireductionist arguments.
What is this theory supposed to achieve? It should clarify antireductionist
arguments, of course. But this answer has to be made more concrete, both
in a substantive and in a methodological respect. Let us begin with the
substantive aspect.

Starting from the plausible assumption that there is more than one kind
of antireductionist arguments, the first step would be to distinguish the
different kinds. In other words, what is needed is a taxonomy of the different
types of antireductionist arguments. The second step in a theory of antire-
ductionist arguments is then a detailed analysis of each of the types of antire-
ductionist arguments. For instance, the following questions have to be
answered for every type of argument: Which premises are used in the
argument? How do these premises lead to the conclusion, i.e., on what is
the argumentative force of the argument based? How strong is the argu-
ment? Where are its limits, that is, how can the argument be invalidated?
This last question is interesting both systematically and historically, since
some at first historically successful antireductionist arguments lost their
persuasiveness in the course of time, and the same may happen to some
contemporary, apparently successful arguments. The third part of a theory
of antireductionist arguments finally consists in a collection and analysis of
common pseudo-arguments. (Note again the parallel of these three parts of a
theory of antireductionist arguments and Aristotle’s conception of logic).

What ways can most usefully be pursued in order to construct a theory
of antireductionist arguments? Generally speaking, the following strategy
seems most promising. First, historical cases of successful antireductionist
arguments must be analyzed, historically as accurately as possible. Second,
the sought theory can be arrived at by reflecting on these cases. The
critical point is, therefore, an appropriate choice of case studies. It is
worthwhile to reflect beforehand upon the conditions these cases must
fulfill to really be case studies for a theory of antireductionist arguments.
I see two necessary and one desirable condition.
At first, the following condition seems to be a necessary condition on the case studies: they must be examples of good antireductionist arguments, since the theory sought after should primarily be a positive theory. This condition is confronted with the difficulty that we do not yet know which arguments indeed are good antireductionist arguments, since it is one of the aims of a theory of antireductionist arguments to inform us about the quality of a particular type of argument. It follows that the condition just suggested has to be replaced by a condition that makes more operational sense. The best replacement seems to be the condition that the cases to be analyzed have to be historically successful. The assessment of an antireductionist argument as historically successful is possible without an evaluation of its argumentative force, and must be based on the historical circumstances. For instance, antireductionist arguments that were formulated by eminent scientists and were, at least for some time, accepted by a majority of the respective scientific community, are good candidates for case studies according to the condition given.

The second necessary condition is of a more technical nature. It states that there must be enough historical data available for the historical reconstruction. This requirement concerns not only the argument itself, but also the scientific background and possibly even some further conditions which contributed to make the argument a convincing argument in its own time.

The fulfillment of a third condition is not necessary although welcome — it is useful to facilitate the work on a theory of antireductionist arguments. According to this requirement, the persuasiveness of the respective argument should have decreased in a later period of history. If this is the case, then it is very likely that we can gain some insight into the limits of the respective type of argument by just studying the historical development. As I said earlier, a theory of antireductionist arguments also comprises the limits of each type of argument.

Happily enough, these three conditions are not so strong that they cannot be satisfied in practice. In the following four cases, at least the first two (necessary) conditions are met.

First case: Aristotle’s cosmology. For Aristotle, the celestial or supra-lunar region is not, in the ontological sense, reducible to the terrestrial or sublunar region. By implication, it is also not reducible in the epistemological sense. The element that constitutes the celestial sphere, the ether, is fundamentally different from the four sublunar elements. Aristotle’s antireductionist argument is mainly contained in his *On the Heavens*, book I, chapters 2 and 3; of course, some of his other works have to play a role in the reconstruction of the background of the argument.
Second case: Hans Driesch’s neovitalism. Driesch developed his well-known neovitalism mainly in his books *Der Vitalismus als Geschichte und als Lehre (The History and Theory of Vitalism)* of 1905 and *The Science and Philosophy of the Organism* of 1908. His main aim was to argue that organisms cannot possibly be conceived of as (mechanical) machines. One of his arguments is based on the idea that a part of a machine cannot contain the whole machine in the sense that, put in a different environment, the part cannot function in the same way as the whole machine. Obviously, Driesch’s machine concept contains as a constitutive element some part-whole asymmetry which makes the parts fundamentally different from the whole. It is this assumption which has — in a sense — been given up in modern biology where every cell contains all the genetic information of the organism. Being one of the founding fathers of experimental embryology, Driesch was influential in biology up to the twenties. Today, neovitalism is almost completely dead, at least in ‘established’ biology.

Third case: Frege’s argument against the reduction of logic to psychology. His argument is mainly contained in the preface of his *Grundgesetze der Arithmetik* of 1893, in his “Logik” of 1897 from the Nachlaß, and in his essay “Der Gedanke” of 1918. The basis of Frege’s arguments is the assertion that there is a substantive difference between the (psychological) process of thinking and what he calls thought (*Gedanke*), the content of thinking. Because of this substantive difference, i.e. because of an ontological antireductionism, the epistemological reduction of logic to psychology is also impossible.

Fourth case: Niels Bohr’s antireductionism inspired by quantum mechanics, concerning the reduction of biology to physics. There is no need to comment on the author and his scientific weight. I am going to sketch Bohr’s argument in the following section.

Further possible case studies concern the relationship between Cartesianism and Newtonianism, the relationship between mechanics and electrodynamics in the 19th century, and the recurrence objection against the reduction of thermodynamics to statistical mechanics.

This brings the more programmatic part of this paper to an end. Now I shall move on to some results of this program, first to the Bohr case study.

3. The Bohr Case Study

For space reasons, I shall here only present a very compressed reconstruction of Bohr’s antireductionist argument (for a somewhat fuller account see Hoyningen-Huene, 1991). In its most elaborate form, the argument was published in Bohr’s paper “Light and Life” of 1933.
Bohr's problem was whether or not the new physics, i.e., classical physics plus quantum mechanics, was sufficient for the explanation of typical biological phenomena. From 1931 on, but not earlier, Bohr's position was that physics and chemistry cannot explain "the peculiar functions characteristic of life", as he puts it (Bohr 1933, p. 9). This position is not identical with epistemological antireductionism, since Bohr talks about explanations and not about derivations. But his position implies epistemological antireductionism since Bohr subscribes to an idea of explanation that is essentially identical with the Hempel-Oppenheim model of explanation, as can be shown by an analysis of the pertinent parts of his paper (see Hoyningen-Huene, forthcoming). Bohr's argument is, therefore, also an argument for epistemological antireductionism.

Bohr's argument can be reconstructed as follows:

**Premise 1:** Whenever complementarity prevails between A and B, it is impossible to explain B by A.

**Premise 2:** Between physics and the typical aspects of life, complementarity prevails.

**Conclusion:** The explanation of typical aspects of life by physics is impossible.

Obviously, this argument is formally valid.

Central to Bohr's argument is the notion of complementarity which has to be discussed first. In "Light and Life", Bohr gives an explanation of this concept with respect to light which comprises the following three aspects:

1. The domain of the complementarity relation consists in different features of one and the same phenomenon, namely in two "equally important features" (Bohr 1933, p. 5). In the case of light, these two features are "the spatial continuity of our picture of light propagation and the atomicity of the light effects" (ibid.).

2. These different features of the phenomenon stand in an "obvious contrast" which "present us with a dilemma of a character hitherto unknown to physics" (Bohr 1933, p. 5). If one assumes the "objective existence of phenomena independent of the means of their observation" (Bohr 1933, p. 7), as one does in classical physics, then the description of light as a wave phenomenon and as a corpuscular phenomenon obviously implies a contradiction. Light can, then, either be a wave or a stream of corpuscles, but not both.

3. The last aspect of complementarity shows how the contradiction is avoided. A "closer analysis in mechanical terms" reveals that the contrasting features show up only under experimental arrangements which are "mutually exclusive" (Bohr 1933, p. 5). There is a physical mechanism that
allows us to avoid the imminent contradiction. In the case of light, it is physically impossible to realize an experimental arrangement in which one and the same beam of light shows both particle and wave features at the same time.

It is a straightforward consequence of the concept of complementarity, namely of its second aspect, that Premise 1 in Bohr’s argument indeed holds. The dilemma arising out of the presence of two seemingly incompatible features indicates that one cannot account for one of the features in terms of the other.

Let us now look how Bohr tries to establish the second premise of his argument. Since Bohr rejects vitalism, he is facing again a “dilemma”, namely, that it is not clear why there should be a difference in the explanatory power of physics with respect to living organisms. Bohr describes the resolution of this dilemma as follows:

In this dilemma it must be kept in mind, however, that the conditions in biological and physical research are not directly comparable, since the necessity of keeping the object of investigation alive imposes a restriction on the former which finds no counterpart in the latter. Thus, we should doubtless kill an animal if we tried to carry the investigation of its organs so far that we could tell the part played by the single atoms in vital functions (Bohr 1933, p. 9).

In other words: the conditions under which an organism is, on the one hand, an object of biology, and, on the other hand, an object of physics, physically exclude each other. It is on this basis that Bohr concludes that complementarity holds between the typical aspects of life and the methods employed by physics, namely “the subdivision necessary for any physical analysis” (Bohr 1933, p. 10). With his argument, Bohr seems, then, to demonstrate that physics cannot explain typical biological phenomena which implies the epistemological irreducibility of biology to physics.

But Bohr’s case for his second premise has an essential weakness. Bohr rests his diagnosis of complementarity mainly on its third aspect, namely, the supposed physical mechanism that avoids a possible contradiction. But a correct diagnosis of complementarity has to establish that all three aspects of complementarity are indeed characteristics of the situation in question. To show that for two features of a phenomenon there is a physical mechanism that excludes the possibility that these two features show up in one and the same situation is by itself theoretically uninteresting. (It is even extremely doubtful that Bohr indeed established the mutual exclusion of physical and biological features; Bohr seems to have realized this later in his life: see Bohr 1963, pp. 198—199). But even if it were true that the physicist’s methods of investigation are physically incompatible with the existence of life, this would imply nothing about reducibility or irreducibility of physics to biology. To take a different example: If it were
impossible for some physical reason to measure the electrical conductivity of a body and its optical reflectivity in one and the same situation, this fact by itself would imply nothing about the theoretical relation between these features.

But isn’t it possible to correct Bohr’s mistake by showing that the two other aspects of complementarity also hold for the relation of physics to typical aspects of life? With respect to the second aspect, this amounts to showing that typically biological features are irreconcilable with those that are the result of a purely physical analysis. But this is just identical with the epistemological irreducibility of these features which we want to establish by means of the diagnosis of complementarity. In other words, it seems that Bohr’s incomplete argument that tries to establish irreducibility by means of complementarity, is circular when completed. The only way to salvage this type of argument would be the discovery of a criterion of complementarity which is independent of its second aspect, but the existence of such a criterion is extremely doubtful.

4. The Theory of Antireductionist Arguments: Further Preliminary Results

The first task of a theory of antireductionist arguments is, as I said earlier, a taxonomy of antireductionist arguments. The arguments I have discussed so far suggest the following preliminary taxonomy (p. 299).

First, those antireductionist arguments that make use of an already established example of antireductionism are distinguished from those which argue from scratch. In the first group, there are arguments which somehow make use of an established example of antireductionism in another domain, and those which make use of an established example of antireductionism of a different kind in the same domain. As far as I can see, the only candidate in the latter class is ontological antireductionism established in that domain. This class splits up in three subclasses. In the first subclass, ontological antireductionism between the domains results from differences in substance; Frege’s antipsychologism is a case in point, perhaps also the Aristotle case. In the second subclass, ontological antireductionism results from differences in the interactions prevailing in otherwise equal substances; Driesch’s antireductionism is a case in point. The third subclass contains all other cases.

In the second major group of arguments which argue from scratch, there is a class of arguments which argue by means of complementarity. This class is possibly empty as is suggested by the Bohr case study (this is why I have put this class in brackets). The remaining class cannot be specified further at this point.
A Theory of Antireductionist Arguments

Antireductionist arguments

by means of an established example of antireductionism

by means of an established example of antireductionism in another domain

by means of ontological antireductionism in the same domain

(by means of complementarity: Bohr case)

from scratch

by means of a difference in substance: Frege case (Aristotle case?)

by means of a difference in interaction: Driesch case

other means

other means

The trick of this taxonomy is, on each level, the introduction of residual classes, i.e., the introduction of the set of all possibilities so far omitted. In this way, a taxonomy is obtained that is indeed complete although one does not really have a clear view of all cases. The danger inherent in this procedure is that a further analysis of the residual classes may disclose cases which suggest a fundamentally different taxonomy (in Hoyningen-Huene 1991 for instance, I have suggested a different taxonomy). This danger can apparently not be avoided.

References


— (forthcoming) "Niels Bohrs Argument für die Nicht-Reduzierbarkeit der Biologie auf Physik". Submitted to Philosophia Naturalis.


