The Perennial Problem of the Reductive Explainability of Phenomenal Consciousness – C. D. Broad on the Explanatory Gap

In: T. Metzinger (ed.) *Neural Correlates of Consciousness – Empirical and Conceptual Questions,* Cambridge MA: MIT-Press 2000, 41-55.

Ansgar Beckermann
University of Bielefeld

Broad's Distinction between Emergent and Mechanically Explainable Properties

At the start of the 20th century the question of whether life could be explained in purely mechanical terms was as hotly debated as the mind-body problem is today. Two factions opposed each other: *Biological mechanists* claimed that the properties characteristic of living organisms (metabolism, perception, goal-directed behavior, procreation, morphogenesis) could be explained mechanistically, in the way the behavior of a clock can be explained by the properties and the arrangement of its cogs, springs, and weights. *Substantial vitalists*, on the other hand, maintained that the explanation envisaged by the mechanists was impossible and that one had to postulate a special nonphysical substance in order to explain life—an entelechy or élan vital. When C. D. Broad developed his theory of emergence in the early 1920s, his aim was to create room for a third position mediating between these two extremes—a position he called *emergent vitalism*.

Broad's first step was to point out that the problem of vitalism is only a special case of a much more general problemthe problem of how the *behavior* of a complex system is related to the *properties and the arrangement of its physical parts*. (According to Broad, living beings differ from nonliving things only in their specific behavior. That is to say, he believed that the property of being alive could be characterized in purely behavioral terms. He strictly distinguished properties of this kind from properties that he called "pure qualities". I will return to this distinction below.)

Regarding this question, there are in principle only two basic types of answers. One can hold the view that the behavior of a complex system cannot be explained by referring exclusively to its 41b
physical parts and their arrangement, but only by the assumption that *S* contains a further, nonphysical component, which is present in all beings that behave in the way characteristic of *S*, and is absent in all other beings. According to Broad, anyone who endorses a theory of this kind is a proponent of a *component theory*. However, one can also hold the opposed view that the behavior of *S can*, at least in principle, be explained by its physical parts and their arrangement. In this case, however, one has, according to Broad, to distinguish two further possibilities. For even if *S*'s behavior can be explained this way, it may be either *mechanistically explainable* or *emergent*. (Broad strictly distinguishes between mechanism and pure mechanism. According to the latter the term "mechanistically explainable" means something like "explainable just by reference to the laws of classical mechanics"; according to former, it means "explainable by reference to all general chemical, physical and dynamical laws" (1925: 46). In the following, "mechanistically explainable" is always meant to have this second, broader meaning.) Mechanistic and Emergent Theories thus concur in denying

> that there need be any peculiar *component* which is present in all things that behave in a certain way and is absent from all things which do not behave in this way. [Both say] that the components may be exactly alike in both cases, and [they] try to explain the difference of behavior wholly in terms of difference of structure." (Broad 1925: 58f.)

However, mechanistic and emergent theories differ fundamentally in their view of the laws that relate the behavior of the components of complex systems to the characteristic behavior of the systems themselves.

> On [the theory of emergence] the characteristic behavior of the whole *could* not, even in theory, be deduced

> from the most complete knowledge of the behavior of its components, taken separately or in other combinations, 42a
> and of their proportions and arrangements in this whole. (Broad 1925: 59)

Which types of macroscopic behavior are to be regarded as emergent in this sense was controversial even in Broad's days. Broad believed that the behavior of nearly all chemical compounds is emergent in the sense explicated. It was his view, for example,

> that, so far as we know at present, the characteristic behavior of Common Salt cannot be deduced from the most complete knowledge of the properties of Sodium in isolation; or of Chlorine in isolation; or of other compounds of Sodium, such as Sodium Sulphate, and of other compounds of Chlorine, such as Silver Chloride. (Broad 1925: 59)

Naturally, mechanists would disagree. For Broad characterizes their theory thus:

> On [the mechanistic theory] the characteristic behavior of the whole is not only completely *determined by* the nature and arrangement of its components; in addition to this it is held that the behavior of the whole could, in theory at least, be *deduced* from a sufficient knowledge of how the components behave in isolation or in other wholes of a simpler kind. (Broad 1925: 59)

Artificial machines are the best examples of complex objects whose behavior can be completely explained in mechanical terms. For instance, we surely do not have any reason to assume that the behavior of a clock is based on a specific non-physical component which is present in clocks and only in clocks. Hence, component theories are quite inappropriate for the explanation of the behavior of these machines. However, we do not have any reason to assume that the behavior of clocks is emergent, either. For, obviously, we can completely deduce its behavior from the specific arrangement of a clock's springs, cogs, weights, and so forth.

Put in a nutshell, the difference between emergent and mechanistic theories can be explained as follows:

> Put in abstract terms the emergent theory asserts that there are certain wholes, composed (say) of constituents *A*, 42b
> *B*, and *C* in a relation *R* to each other; that all wholes composed of constituents of the same kind as *A*, *B*, and *C* in relations of the same kind as *R* have certain characteristic properties; that *A*, *B*, and *C* are capable of occurring in other kinds of complex where the relation is not of the same kind as *R*; and that the characteristic properties of the whole *R(A,B,C)* cannot, even in theory, be deduced from the most complete knowledge of the properties of *A*, *B*, and *C* in isolation or in other wholes which are not of the form *R(A,B,C)*. The mechanistic theory rejects the last clause of this assertion. (Broad 1925: 61)

Broad here stresses two points: First, regardless of whether the characteristic behavior *B* of a class of systems is mechanistically explainable or emergent, *B* nomologically depends on the corresponding microstructures. That is to say, if a system *S* consists of the parts $C_1, \ldots, C_n$ arranged in manner *R*—for short: if *S* has microstructure $[C_1, \ldots, C_n; R]$, then the sentence "All systems with microstructure $[C_1, \ldots, C_n; R]$ behave in manner *B*" is a true law of nature—regardless of whether *B* is emergent or mechanistically explainable. Obviously, this is the reason for Broad's view that emergent as well as mechanistically explainable properties can be explained by reference to the microstructure of the respective system. However, Broad here employs a comparatively weak notion of explanation.

Second, mechanistically explainable behavior differs from emergent behavior in that the former can, at least in principle, be deduced "from the most complete knowledge of the properties of [the components $C_1, \ldots, C_n$] in isolation or in other wholes" while this cannot be done for the latter.

Broad's concepts of mechanistic explainability and emergence can thus be summarised as follows:

(ME) The characteristic behavior *B* of a complex system *S* with the microstructure $[C_1, \ldots,$
$C_n; R]$ is *mechanistically explainable* if and only if *B* can (at least in principle) be de- 43a
duced from the most complete knowledge of all properties that the components $C_1, \ldots, C_n$ have either in isolation or within other arrangements.

(E) The characteristic behavior *B* of a complex system *S* with the microstructure $[C_1, \ldots, C_n; R]$ is *emergent* if and only if the following is true:

(a) The statement “All systems with microstructure [C_1, …, C_n; R] behave in manner *B*” is a true law of nature, but

(b) *B* cannot (even in principle) be deduced from the most complete knowledge of all properties that the components C_1, ..., C_n have either in isolation or within other arrangements.

The general upshot of these definitions seems to be clear enough. But why does Broad use the complicated clause “from the most complete knowledge of the properties of [the components C_1, …, C_n] in isolation or in other wholes”?

To begin with, Broad evidently saw that the notion of an emergent property would, for quite trivial reasons, be empty if one were permitted to use all the properties of the components in the “deduction” of the behavior *B*. Some twenty years after the first publication of *The Mind and Its Place in Nature*, Hempel and Oppenheim, referring to a remark by Grelling, phrased this problem thus:

> If a characteristic of a whole is counted as emergent simply if its occurrence cannot be inferred from a knowledge of all the properties of its parts, then, as Grelling has pointed out, no whole can have any emergent characteristics. Thus … the properties of hydrogen include that of forming, if suitably combined with oxygen, a compound which is liquid, transparent, etc. Hence the liquidity, transparency, etc. of water *can* be inferred from certain properties of its chemical constituents. (Hempel/Oppenheim 1948: 260)

In order to avoid rendering the concept of emergence vacuous, inferences of this kind must be blocked. Broad’s formula serves precisely this purpose, since it is obviously designed to guarantee that we cannot have recourse to properties like those mentioned by Hempel and Oppen- 43b
heim when we attempt to deduce the characteristic behavior *B* of a complex system from the properties of its parts and its structure. However, the question remains whether this purpose could have been accomplished with a simpler and more lucid formulation. This much seems clear: It is crucial that in our attempts to deduce some behavior *B* of a complex object from the properties of its parts and their spatial relations, we are not allowed to use ad hoc properties such as the property that certain components, if arranged in a specific way, form a complex object which behaves in manner *B*. The question, therefore, is how we can guarantee this result without at the same time excluding properties which we may legitimately invoke in such an attempt.

An answer to this question can be found if we consider which laws we may use in deductions of this type. For here we encounter a related possibility of trivialising the concept of emergence. If we were allowed to employ the law mentioned above, i.e. the law “All systems with microstructure [C_1, …, C_n; R] behave in manner *B*”, there would be no emergent behavior. Hence, Hempel and Oppenheim could have formulated their point just as well in this way: It is a true law of nature that, if suitably combined with oxygen, hydrogen forms a compound which is liquid, transparent, and so on. Hence the liquidity, transparency, and such of water *can* be derived by means of the laws of nature. Clearly, Broad was aware of both possible ways of trivialising the concept of an emergent property (see, e.g., Broad 1925: 65ff.).

Broad, therefore, must also rule out recourse to laws of this type. That this is something he actually sought to do can be seen from the following passage discussing the properties of clocks:

> We know perfectly well that the behavior of a clock can be deduced from the particular arrangement of springs, wheels, pendulum, etc., in it, and from *general laws of mechanics and physics which apply just as much to ma-* 44a
> *terial systems which are not clocks*. (Broad 1925: 60 – italics added)

Obviously, Broad held that if we attempt to deduce some behavior *B* of a complex object from the properties and arrangement of its parts, we may use only *general laws* that are valid for the parts of a complex system independently of the specific configurations of these parts. However, this constraint provides a way to rule out recourse to ad hoc properties as well. Hence, the most straightforward answer to the question “Which properties of a system’s parts may we refer to in such a deduction?” is apparently this: “to those properties which are mentioned in these general laws of nature". I should therefore like to suggest that we replace Broad's clause with the formula "if *B* can be deduced, by means of the *general* laws of nature that are true of the components C_1, ..., C_n, from the properties of the components mentioned in these laws". Taken to its logical conclusion, this improved version of Broad's formula renders superfluous any reference to admissible properties; if we specify which laws can figure in the derivations in question, we have implicitly determined which properties may play a role in these derivations.

Even after this point has been clarified, however, the question remains why, according to Broad, we need to know not only how the components of a system behave "in isolation" but also how they behave "in other wholes". As we have already seen, Broad thought that mechanistically explainable behavior differs from emergent behavior in that the former can be deduced, by means of the general laws of nature which are true of the components C_1, …, C_n, from the properties of the components mentioned in these laws, whereas the latter cannot. But this provokes the further question of how we can determine which laws are general in the sense required. Broad's own answer to this question has two parts: First, we have to observe how the parts behave in isolation, and second, we have to investigate how they behave in "other" systems.

Why do we have to do both? Broad was quite obviously thinking of the dynamic behavior of 44b
systems that are subject to a number of different forces (see Broad 1925: 62, 63f). If we want to find out whether the law that is crucial here, the second Newtonian law $F = m \cdot a$, applies in this case, we have to begin by investigating the behavior of objects that are subject to only *one* force. But if we wish to know how an object behaves generally—that is, how it behaves if more than one force acts on it simultaneously— the knowledge of this law is not enough. We also have to know the law that governs the interaction of the various forces: the law of the vector addition of forces. According to Broad, we always need these two types of laws: (a) laws that state how individual factors separately influence the behavior of an object and (b) laws that state what behavior results if different factors simultaneously act on an object. Laws of the second type Broad terms "laws of composition". Moreover, he emphatically stresses their indispensability:

> It is clear that in *no* case could the behavior of a whole composed of certain constituents be predicted *merely* from a knowledge of the properties of these constituents, taken separately, and of their proportions and arrangements in the particular complex under consideration. Whenever this *seems* to be possible it is because we are using a suppressed premise which is so familiar that it has escaped our notice. The suppressed premise is the fact that we have examined other complexes in the past and have noted their behavior; that we have found a general law connecting the behavior of these wholes with that which their constituents would show in isolation; and that we are assuming that this law of composition will hold also of the particular complex whole at present under consideration. (Broad 1925: 63)

However, it is not completely clear what kind of law Broad is alluding to here. The way Broad speaks in the passage just quoted, it seems as if laws of composition are meant to relate the behavior of a *system* to the behavior of its *parts*. (Only the phrase "in isolation" is puzzling in this interpretation.) In this case, laws of composition would have the status of bridge principles re- 45a
lating the level of the parts to the level of the whole. Yet directly after this passage Broad returns to the example of the explanation of the dynamical behavior of objects that are subject to a number of forces:

> For purely dynamical transactions this assumption is pretty well justified, because we have found a simple law of composition and have verified it very fully for wholes of very different composition, complexity, and internal structure. It is therefore not particularly rash to expect to predict the dynamical behavior of any material complex under the action of any set of forces, however much it may differ in the details of its structure and parts from those complexes for which the assumed law of composition has actually been verified. (Broad 1925: 63f.)

Obviously, the law of composition he refers to in this passage is the law of the vector addition of forces already mentioned. (see Broad 1925: 62).

This law, however, does not state how the behavior of a whole arises from the behavior of its parts, but how the parts of a whole behave if they are subject to a number of forces. Thus it might be more apt to call laws of this type "laws of interaction".

Fortunately, we do not have to decide on one reading, for Broad seems to be right in either case.

On the one hand, we of course need laws of interaction since we cannot deduce the behavior of a system from the properties of its parts and their arrangement if we do not know how the parts themselves move if they are arranged in this particular way. On the other hand, we also need laws of composition or bridge principles since we cannot deduce the behavior of a system from the behavior of its parts if we do not know how the behavior of the parts is related to the behavior of the whole. Thus, in order to deduce the behavior of a system from the properties of its parts and their arrangement, we actually need three types of laws:

1. Simple laws, which state how each part of the system *S* behaves if only a single factor acts on it
2. Laws of interaction, which state how the parts of *S* behave if a number of factors simultane- 45b
ously act on them
3. Laws of composition or bridge principles, which state how *S* behaves as a whole if its parts behave in a specific way.

The fact that any attempt to explain the behaviour of a system *S* by reference to the properties of its parts and their arrangement requires three types of laws is also pointed out by Hüttemann and Terzidis (in press). The indispensability of laws of composition is stressed in McLaughlin (1992).

It should again be noted, that all these laws must be general laws or must follow from general laws to be usable in what Broad calls the deduction of the characteristic behavior of a system from the properties and the arrangement of its parts. Put precisely, what Broad's definitions come to, therefore, is the following:

(ME′) The characteristic behavior *B* of a complex system *S* with the microstructure [C_1, …, C_n; *R*] is *mechanistically explainable* if and only if the following is true:

- (a) The way the components C_1, ..., C_n behave when arranged in manner *R* can be accounted for by the general simple laws and by the general laws of interaction holding for objects of the kind C_1, ..., C_n; and
- (b) there is a general law of composition to the effect that *S* exhibits behavior *B* if the components C_1, ..., C_n behave in the way they do.

(E′) The characteristic behavior *B* of a complex system *S* with the microstructure [C_1, …, C_n; *R*] is *emergent* if and only if the following is true:

- (a) The statement "All systems with microstructure [C_1, …, C_n; *R*] behave in manner *B*" is a true law of nature, but
- (b_1) the way the components C_1, ..., C_n behave when arranged in manner *R* cannot be accounted for by the general simple laws and by the general laws of interaction holding for objects of the kind C_1, ..., C_n; or
- (b_2) there is no general law of composition to the effect that *S* exhibits behavior *B* if 46a
the components C_1, ..., C_n behave in the way they do.

Two points should be highlighted here. The first concerns the question of why Broad did not provide a positive example for a law of composition that would render the behavior of a given system mechanistically explainable. My guess is that the laws Broad had in mind are so mundane that he felt no need to mention them. For instance, the following seems to be trivially true of spatial movement:

(P1) If we know how all the components of a complex system move, we also know how the system itself moves.

Think, for example, of a disk whose components all revolve with the same angular velocity and in the same direction around the disk's centre. Then it seems quite clear that the disc as a whole spins round its center. Indeed, given the movements of its components, it seems inconceivable that the disk itself moves in any other fashion. (A quite similar point could be made with regard to clocks.)

In much the same way the volume of a system seems to be determined by the places its components occupy, and the shape of a system seems to be determined by the relative positions of its components. The bridge principles Broad must have had in mind thus do not just have the character of very general laws of nature; rather, since we simply cannot conceive that they could be false, they seem to have the status of a priori truths.

This brings us to the second point. Broad himself emphasizes that the status of the law "All systems with microstructure [C_1, …, C_n; *R*] behave in manner *B*" varies remarkably depending upon whether the behavior in question can be explained in a mechanical way or only in an emergent one. With respect to emergent behavior, this law has the status of a unique and ultimate law. It is (a) not a special case that can be arrived at by substituting determinate values for 46b
determinable variables in general laws. It is (b) not a law that can be generated by combining two or more general laws. And, perhaps most important, (c) the law in question is a law that can be discovered *only* by studying samples of systems with the microstructure [C_1, …, C_n; *R*] and can be extended inductively *only* to other systems with the same microstructure. Regarding silver chloride Broad writes:

> … the law connecting the properties of silver-chloride with those of silver and of chlorine and with the structure of the compound is, so far as we know, an *unique* and *ultimate* law. By this I mean (*a*) that it is not a special case which arises through substituting certain determinate values for determinable variables in a general law which connects the properties of *any* chemical compound with those of its separate elements and with its structure. And (*b*) that it is not a special case which arises by combining two more general laws, one of which connects the properties of *any* silver-compound with those of elementary silver, whilst the other connects the properties of *any* chlorine-compound with those of elementary chlorine. So far as we know there are no such laws. It is (*c*) a law which could have been discovered *only* by studying samples of silver-chloride itself, and which can be extended inductively *only* to other samples of the same substance. (Broad 1925: 64f.)

If, on the other hand, we are concerned with mechanistically explainable behavior, things are very different:

> In order to predict the behavior of a clock a man need never have seen a clock in his life. Provided he is told how it is constructed, and that he has learnt from the study of *other* material systems the general rules about motion and about the mechanical properties of springs and of rigid bodies, he can foretell exactly how a system constructed like a clock must behave. (Broad 1925: 65)

If the behavior *B* of a system *S* can be explained mechanistically, we are in a position to know that *S*— like all systems with the microstructure [C_1, …, C_n; *R*] —behaves in manner *B* without having investigated a single system with this particular microstructure. For in this case 47a
the behavior of the system follows directly from the general laws of nature that apply to the components C_1, …, C_n. (These general laws of nature comprise the simple laws as well as the laws of interaction and the laws of composition.) This implies that in the case of mechanistically explainable behavior, it is in a certain sense *inconceivable* that a system has the microstructure [C_1, …, C_n; *R*], but does not exhibit behavior *B*. For if it follows from the general laws of nature that all systems with that microstructure behave in manner *B*, then it is impossible—at least *relative to these laws of nature*—that a system possesses the microstructure [C_1, …, C_n; *R*], but does not behave in manner *B*. Thus it is a conclusive *test* for the behavior of a system *S* being mechanistically explainable that it can be predicted before it first occurrs or, in other words, that—relative to the laws of nature—it is inconceivable that it does not occur, given the microstructure of *S*.

Chemical and secondary qualities as candidates for emergent properties

I shall now turn to the following questions: Which properties did Broad take to be emergent and what were his reasons for doing so? We already know what his reasons were with respect to the characteristic behavior of chemical compounds. According to Broad, there are no appropriate laws of composition:

> The example of chemical compounds shows us that we have no right to expect that the same simple law of composition will hold for chemical as for dynamical transactions…. It would of course (on any view) be useless merely to study silver in isolation and chlorine in isolation; for that would tell us nothing about the law of their conjoint action. This would be equally true even if a mechanistic explanation of the chemical behavior of compounds were possible. The essential point is that it would also be useless to study chemical compounds in gen- 47b
> eral and to compare their properties with those of their elements in the hope of discovering a *general* law of composition by which the properties of *any* chemical compound could be foretold when the properties of its separate elements were known. So far as we know, there is no general law of this kind…. No doubt the properties of silver-chloride are completely *determined* by those of silver and of chlorine; in the sense that whenever you have a whole composed of these two elements in certain proportions and relations you have something with the characteristic properties of silver-chloride, and that nothing has these properties except a whole composed in this way. But the law connecting the properties of silver-chloride with those of silver and of chlorine and with the structure of the compound is, so far as we know, an *unique* and *ultimate* law. (Broad 1925: 64f.)

Much the same holds, Broad thought, for the characteristic features of living beings:

> … it is obviously possible that, just as the characteristic behavior of a [chemical] compound could not be predicted from any amount of knowledge of the properties of its elements in isolation or of the properties of other [chemical] compounds, so the properties of a [compound that is made up of chemical compounds, i.e., a living body] could not be predicted from any amount of knowledge about the properties of its [chemical] constituents taken separately or in other surroundings …; so the only way to find out the characteristic behavior of living bodies may be to study living bodies as such. (Broad 1925: 67)

Nowadays, Broad's treatment of both examples seems outdated. We now know that the electrical conductivity of metals is due to the fact that only a few electrons are located in their outer shell, and these electrons are capable of relatively free movement. We do also know that the metal sodium combines with chlorine because chlorine atoms can complete their outer electron shell with the electrons given off by the sodium atoms. In this process sodium and chlorine ions are created; they exert strong electrical forces on each other and, therefore, form a lattice structure. We know, moreover, that sodium chloride is water-soluble because water molecules can—
due to their dipole structure—pull the sodium ions and the chlorine ions from their places within 48a
the lattice. And we now also know much about the chemical processes on which the breathing, digestion and reproduction of living beings are based.

However, it should be noted that Broad did not rule out any of these developments. For example, he writes with regard to the process of breathing:

> … since [the process of breathing] is a movement and since the characteristic movements of some complex wholes (*e.g.*, clocks) *can* be predicted from a knowledge of their structure and of other complex wholes which are not clocks, it cannot be positively *proved* that breathing is an "ultimate characteristic" or that its causation is emergent and not mechanistic. Within the physical realm it always remains logically possible that the appearance of emergent laws is due to our imperfect knowledge of microscopic structure and to our mathematical incompetence. (Broad 1925: 81)

Yet Broad immediately adds: "But this method of avoiding emergent laws is not logically possible for trans-physical processes …" (Broad 1925: 81). But what does Broad mean by "trans-physical processes"? And why does he believe that such processes could never be explained mechanistically? In order to understand this, we have to grasp Broad's concept of a pure quality. For Broad considers those laws "to be trans-physical" which relate a system's microstructure to its pure qualities—that is to say, all laws of the form "All systems with microstructure [C_1, …, C_n; R] have pure quality F" (see Broad 1925: 52).

Broad's official definition of "pure quality" is this:

> By calling [qualities such as red, hot, etc.] "pure qualities", I mean that, when we say "This is red", "This is hot" and so on, it is no part of the meaning of our predicate that "this" stands in such and such relation to something else. It is *logically* possible that this should be red even though "this" were the only thing in the world …. (Broad 1925: 52)

This definition, however, is rather puzzling, since for Broad pure qualities seem to be exactly 48b
those properties of complex objects which traditionally are called secondary qualities—tem-

perature, color, taste, and smell (see Broad 1925: 46ff., 79f.). Yet secondary qualities are usually characterized as nothing but powers to cause certain sensations in us. So how can an object possess pure qualities unless there are creatures perceiving it? But let us return to the question of why Broad believed that pure qualities are necessarily emergent.

As we have already seen, Broad held the view that “so far as we know at present” most of the *behavior* of chemical compounds is emergent. Yet he thought that scientific progress might prove this view wrong. However, chemical compounds are characterized not only by their specific behavior but also through their pure qualities. May it turn out that these pure qualities, too, are not emergent, but mechanistically explainable? Broad’s answer to this question was an unequivocal no. (For the following, see Broad 1925: 71f.).

Consider ammonia, for instance. Ammonia is a gas whose molecules consist of three hydrogen atoms and one nitrogen atom. It is readily soluble in water and has a pungent smell. Possibly (according to Broad) we shall one day be able to explain the water solubility of ammonia as well as its other characteristic behavior by reference to the properties of its components and their arrangement. However, this is not true of its smell. Broad thought it to be *theoretically* impossible to explain this smell in a mechanical way. Why? Broad’s answer was this: Not even a mathematical archangel—that is to say, a creature who knows all general laws of nature and is able to execute the most complex mathematical calculations in split seconds—could predict what the compound from three hydrogen atoms and one nitrogen atom would smell like.

> [Even a mathematical archangel] would be totally unable to predict that a substance with [the microscopic 49a
> structure of ammonia] must smell as ammonia does when it gets into the human nose. The utmost that he could predict on this subject would be that certain changes would take place in the mucous membrane, the olfactory nerves and so on. But he could not possibly know that these changes would be accompanied by the appearance of a smell in general or the peculiar smell of ammonia in particular, unless someone told him so or he had smelled it for himself. If the existence of the so-called “secondary qualities” … depends on the microscopic movements and arrangements of material particles which do not have these qualities themselves, then the laws of this dependence are certainly of the emergent type. (Broad 1925: 71f.)

Why would even a mathematical archangel be limited in this way? Considering what we have said so far, the reason can only be that it simply does not follow from the general laws applicable to hydrogen atoms and nitrogen atoms that a compound of three hydrogen atoms and one nitrogen atoms smells in the way characteristic of ammonia. From these laws (and the laws of neurophysiology) it follows at best that in the cells of the olfactory nerves and in the brain of a person whose mucous membrane is hit by ammonia molecules, certain electrochemical changes take place. What does not follow is that these changes are accompanied by certain olfactory sensations. In other words, the law which states that certain changes in a person’s nervous system lead to such olfactory sensations is not deducible from the general laws of nature. It is an emergent law or, as Broad also calls it, a “trans-ordinal law”—a law that relates a system’s microstructure to its nondeducible properties.

Broad’s argument for the thesis that pure qualities are necessarily emergent can thus be summarized as follows: Pure qualities are secondary qualities, and it is a characteristic feature of secondary qualities that they cause certain sensations in us. But it does not follow from the general laws of nature that a system *S*, which has a certain microstructure, will cause a certain sensation
in us. At most it follows from these laws that certain changes will be caused in our nervous sys- 49b
tem by light reflected from *S* or by molecules emitted into the air by *S*. The pivotal point of this argument is Broad’s thesis that the laws of composition which relate certain processes in our CNS to our sensations are unique and ultimate laws that cannot be derived from the general laws of nature. Broad’s main reason for asserting the emergent character of pure qualities is thus his assumption that sensations cannot be deduced from occurrences in a person’s CNS. We must therefore examine the arguments Broad devises to sustain this crucial assumption.

A central reason for this assumption is certainly that Broad believed that sensations—like all other mental states—cannot be analyzed in behavioristic terms. (For the following, see especially Broad 1925: 612-624. Stephan’s (1993) analysis of Broad’s arguments is similar to my

own.)
In Broad's view, this is the main difference between the mind-body problem and the problem of vitalism.

> The one and only kind of evidence that we ever have for believing that a thing is alive is that it behaves in certain characteristic ways. *E.g.,* it moves spontaneously, eats, drinks, digests, grows, reproduces, and so on. Now all these are just actions of one body on other bodies. There seems to be no reason whatever to suppose that "being alive" means any more than exhibiting these various forms of bodily behavior.… But the position about consciousness, certainly seems to be very different. It is perfectly true that an essential part of our evidence for believing that anything but ourselves has a mind and is having such and such experiences is that it performs certain characteristic bodily movements in certain situations…. But it is plain that our observation of the behavior of external bodies is not our only or our primary ground for asserting the existence of minds and mental processes. And it seems to me equally plain that by "having a mind" we do not mean simply "behaving in0 such and such ways". (Broad 1925: 612f)

For Broad the falsity of behaviorism follows mainly from two considerations: (1) My self-ascriptions of mental states are not based on observable behavior. Even if my body behaved in 50a
one characteristic way when I see a chair and in another when I hear a bell

> it is perfectly certain that this is not my ground for saying that I see a chair or hear a bell. I often know without the least doubt that I am having the experience called "seeing a chair" when I am altogether uncertain whether my body is acting in any characteristic way. And again I distinguish with perfect ease between the experience called "seeing a chair" and the experience called "hearing a bell" when I am quite doubtful whether my bodily behavior, if any, on the two occasions has been alike or different. (Broad 1925: 614)

(2) Every good actor can behave exactly like someone who feels pain or great joy. Hence one cannot conclude from the fact that someone behaves in a certain way that he really has certain sensations or perceptions. More generally, this means that even if some creature *A* always acts exactly like someone who has real sensations, we can always ask: Does *A* really have sensations or does he only behave as if he had them? With respect to intelligent behavior, Broad formulates this point as follows:

> However completely the behavior of an external body answers to the behavioristic tests for intelligence, it always remains a perfectly sensible question to ask: "Has it really got a mind, or is it merely an automaton?" It is quite true … that, the more nearly a body answers to the behavioristic tests for intelligence, the harder it is for us in practice to contemplate the possibility of its having no mind. Still, the question: "Has it a mind?" is never silly in the sense that it is meaningless.… it is not like asking whether a rich man may have no wealth. (Broad 1925: 614)

So Broad was apparently an early proponent of the argument from absent qualia and believed in the possibility of philosophical zombies. It should be noted, however, that even if Broad's criticism of behaviorism succeeds, this does not by itself show that every law of the form "If the
neurological process *N* takes place in person *A*'s CNS, then *A* has the sensation *E*" has to be an 50b
emergent law. Why should not at least some laws that relate the microstructure of a system to one of its properties that is *not* behaviorally analyzable fail to be emergent?
A part of the answer to this question resides, I think, in the fact that Broad seems to have believed that laws such as

(1) "Whenever C-fibers fire in a person's CNS, this person feels pain"

differ in their status from general bridge principles like:

(P1) "If one knows how all components of a complex system *S* move, one also knows how *S* itself moves;
(P2) if one knows the places of all the components of *S*, one knows what volume is occupied by *S*;
(P3) If one knows the relative positions of all the components of *S*, one knows the shape of *S*".

For, in a sense, we simply cannot conceive that these principles could be false while, by contrast, we can very well imagine law (1) to be false. Hence (1) does not have the status that would be required for it to play the role of a general bridge principle.

Broad gives another argument, which shows that he anticipated not only the explanatory gap but also the knowledge argument:

> We have no difficulty in conceiving and adequately describing determinate possible motions which we have never witnessed and which we never shall witness.… But we could not possibly have formed the concept of such a colour as blue or such a shade as sky-blue unless we had perceived instances of it, no matter how much we had reflected on the concept of Colour in general or on the instances of other colours and shades which we *had* seen. It follows that, even when we know that a certain *kind* of secondary quality … pervades … a region when and only when such and such a *kind* of microscopic event … is going on within the region, we still could not possibly predict that such and such a determinate event of the kind … would be connected with such and 51a
> such a determinate shade of colour … The trans-physical laws are then *necessarily* of the emergent type. (Broad 1925: 80)

According to Broad, it is thus crucial that we can form the concept of a certain sensation only after we ourselves have experienced this sensation for the first time. (Although Broad here speaks about secondary qualities, we have seen earlier that the concern really lies with the sensations which are caused by secondary qualities.) However, if it is the case that we can form the concept of a sensation only after experiencing it, the occurrence of an experience cannot be predicted before it has been experienced at least once. From this alone it follows that sensations are emergent, since for reducible properties—by contrast—it is characteristic that they can be predicted before their first occurrence.

There is one point I should like to add here. Broad obviously thought that sensations are emergent because the laws of composition which connect neural events in the CNS of a person to the person's sensations are unique and ultimate, and therefore lack the status of general bridge laws. At first sight, he seems to have held very much the same view with regard to the emergence of the behavior of chemical compounds. Broad wrote, "The example of chemical compounds shows us that we have no right to expect that the same simple law of composition will hold for chemical as for dynamical transactions" (Broad 1925: 64).

The comparison with the law of dynamical transactions in mechanics, however, makes it quite clear that what is at issue in chemistry is not laws of composition but laws of interaction. According to Broad, the emergent character of the behavior of chemical compounds results from the fact that there are no general laws which tell us how *atoms* behave in all possible arrangements. With regard to definition (E′) the emergent character of the behavior of chemical compounds is thus due to condition (b_1). In the case of sensations, things seem to be different, even 51b
though Broad is not very clear on this point.[1] But he seems to have held that even if there were general laws of interaction which told us everything we wanted to know about the neural events going on in the CNS, we could not predict the sensations connected with these events, because of the lack of suitable laws of composition or bridge principles. Thus, the emergent character of sensations is due to condition (b_2) rather than (b_1).

Broad's Account of the Explanatory Gap

It should have become clear from this account of Broad's position on the emergent character of sensations that his thoughts contain all the essential ingredients of the explanatory gap argument. A direct comparison with Levine's and Chalmer's arguments would make this even clearer. Since I cannot do both, I restrict myself here to the work of Levine. Let me briefly rehearse his main line of argument. Consider the two statements: (1) Pain in humans is identical to C-fiber firing and (2) Temperature in ideal gases is identical to the mean kinetic energy of their molecules.

According to Levine, there is a fundamental difference between these two statements—while the second is "fully explanatory", whereas the first is not. That is to say, on the one hand it is *inconceivable* in a certain epistemological sense that in a gas the mean kinetic energy of the molecules has a certain numerical value (say, $6.21 \cdot 10^{-21}$ Joule), but that the gas does not have the corresponding temperature of 300 K. On the other hand, it seems perfectly *conceivable* that I do not feel any pain although my C-fibers are firing. What is the reason for this difference?

Levine's answer is this. If we were asked what we mean by the term "temperature", we would answer: (2′) Temperature is the property of bodies that causes certain sensations of warmth and coldness in us, and that causes the mercury column of thermometers that come into contact with 52a
these bodies to rise or fall, and that causes certain chemical reactions, and so forth.

In other words: We would characterize temperature by its causal role alone. However, this would not be a sufficient answer to the question if there were not a second point:

> … our knowledge of chemistry and physics makes intelligible how it is that something like the motion of molecules could play the causal role we associate with heat. Furthermore, antecedent to our discovery of the essential nature of heat, its causal role, captured in statements like (2′), exhausts our notion of it. Once we understand how this causal role is carried out there is nothing more we need to understand. (Levine 1983: 357)

Thus the explanatory character of statement (2) is due to *two* facts: (1) Our concept of temperature is exhausted by the causal role of temperature; (2) Physics can make it intelligible that the mean kinetic energy of the molecules of a gas plays exactly this causal role.

Now it is plain that we also associate a causal role with the term "pain". Pain is caused by injury to tissue, it causes us to scream or whimper, and it causes us to wish to be rid of it as soon as possible. Levine neither disputes this nor denies that the identification of pain with the firing of C-fibers explains the mechanism that underpins this causal role. Nevertheless, there is a crucial difference between the two cases:

> However, there is more to our concept of pain than its causal role, there is its qualitative character, how it feels; and what is left unexplained by the discovery of C-fiber firing is *why pain should feel the way it does*! For there seems to be nothing about C-fiber firing which makes it naturally "fit" the phenomenal properties of pain, any more than it would fit some other set of phenomenal properties. Unlike its functional role, the identification of the qualitative side of pain with C-fiber firing … leaves the connection between it and what we identify it with completely mysterious. One might say, it makes the way pain feels into merely a brute fact. (Levine 1983: 357)

Thus, Levine's first reason for his thesis that (1) is not fully explanatory is the following: Pain is 52b
only partly characterized by its causal role. For it is also a characteristic feature of pain that it has a certain qualitative character, namely, that it is painful.

However, this argument on its own does not suffice to demonstrate that (1) is not fully explanatory. For it at least could be the case that it follows from the laws of neurobiology that C-fiber firings also possess this characteristic feature. This is precisely what Levine has to deny. Hence, Levine's second thesis is: It does not follow from the laws of neurobiology that C-fiber firings possess the qualitative character of pain, that is, that C-fiber firings feel painful.

With regard to the latter claim one could ask, however, whether this might not be due to the fact that, at the moment, neurobiological research is not sufficiently advanced. Is it really completely impossible that neurobiology will one day tell us that C-fiber firings in a certain sense must, after all, be connected to the qualitative character of pain?

When answering this question in his article "On Leaving Out What It's Like", Levine begins by stressing that every reduction has to lead to an *explanation* of the phenomenon reduced, and that if this explanation is successful, it is indeed impossible—in an epistemic sense—to conceive of the occurrence of the explanans without the explanandum.

> The basic idea is that a reduction should explain what is reduced, and the way we tell whether this has been accomplished is to see whether the phenomenon to be reduced is epistemologically necessitated by the reducing phenomenon, i.e., whether we can see why, given the facts cited in the reduction, things must be the way they seem on the surface (Levine 1993: 129).

Let us try to unpack what Levine has in mind here by means of the example of the macroproperty of being liquid. In general, liquids differ from gases in that their volume is (almost) incom- 53a
pressible. They differ from solids in that their shape is changeable and moulds itself to the receptacle holding them. This provides us with an—albeit incomplete—list of the features that characterize the property of being liquid. Can we now derive the fact that water is liquid at a temperature of 20° C from the properties of its molecules? Does it follow from the general laws

of nature which apply to H_2O molecules that at a temperature of 20° C (and normal pressure) water has all the features which are characteristic of the property of being liquid?
From the laws of nature it follows, first,[2] that the mean distance of H_2O molecules from each other can be reduced further only through great pressure because of the repulsive forces between the molecules. Second, it follows from the laws of nature that the attractive forces between the molecules are not strong enough to fix the molecules in their relative positions. Therefore, the molecules can freely roll over each other. If all molecules are subject to the same force, each molecule will move to the place from which it cannot move any further.
However, this alone does not show that *water* at a temperature of 20° C has all the features which are characteristic of the property of being liquid. For up to now we know only how *the individual H_2O molecules* behave at this temperature. That is to say, in addition we need *bridge principles* (see Levine 1993: 131) which state how the behavior of the liquid as a whole is related to the behavior of the individual molecules. These principles are obviously the following:

(P4) If the mean distance between the molecules of some substance can be reduced only by great pressure, then the volume of that substance can be reduced only by great pressure.

(P5) If the molecules of some substance can freely roll over one another, then the shape of this substance is flexible and moulds itself to the shape of the receptacle in which the substance is placed.

This leads us to the following answer to the question why—once this explanation has been 53b
given—we cannot conceive that water is *not* liquid at a temperature of 20° C. The first reason is simply that it follows from the general laws of nature that the mean distance between H_2O molecules at a temperature of 20° C can be reduced further only by great pressure and that the attractive forces between the molecules are not strong enough to fix them in their relative positions. The second reason, which is bound up with the special status of the bridge principles (P4) and (P5), is just as important. For obviously this status is responsible for the fact that we *cannot* imagine that the mean distance between the molecules of some substance can be reduced only by great pressure, but that the volume of this substance already decreases under little pressure, or that the molecules of some substance may freely roll over one another but the shape of this substance is inflexible, and therefore does not mould itself to the shape of the receptacle in which it is placed.
According to Levine, things are very different if we consider the relation between pain and C-fiber firings. Even if we know to the smallest detail what neurophysiological processes occur in a person's brain, it is still conceivable (according to Levine) that the person in whose brain these processes occur does not feel any pain. What is the reason for this difference?
If we consider a detailed analysis of the explanation of why it is inconceivable that water is not liquid at a temperature of 20° C, three points stand out:

1. *All* of the characteristic features of the property's being liquid consist in the fact that under certain conditions liquid substances *behave* in a specific way.
2. It follows from the general laws of nature that at a temperature of 20° C certain attractive and repulsive forces obtain between H_2O molecules.
3. There are general bridge principles which state that a substance between whose molecules 54a
these forces obtain displays exactly such behavior as is characteristic of the property of being liquid.

Now what about the alleged explanation of pain through the firing of C-fibers? Obviously, when we think about the first point, we immediately encounter a clear difference:

1. Our concept of pain is not exhausted by its causal role, and pain is not just characterized by a certain kind of behavior. Rather, our concept of pain includes a qualitative aspect—what it feels like to be in pain.

However, this point is not the heart of the matter. For pain could still be explained by the firing of C-fibers *if only* there were bridge principles to the effect that the firing of C-fibers feels the way that is characteristic of pain. Therefore the following two points are really crucial here:

2. From the laws of neurobiology it follows *only* under what conditions which neurons fire with what frequency.
3. There are no general bridge principles that connect the firing of neurons with certain qualitative experiences.

Hence Levine's explanatory gap argument, as well as Broad's argument for the emergent character of secondary qualities and sensations, are based on the central thesis:

(T1) Laws like (1), which connect neural processes with sensations, do not have the status of general bridge laws.

For Levine as well as for Broad, the emergent character of sensations thus results from condition (b_2) rather than from condition (b_1) of definition (E′). That is, the problem is not that there are no general laws of *interaction* which tell us how neurons fire if they are connected in the way
they are connected in our CNS. The real problem is that the laws which tell us what goes on in 54b
our mind if certain neuron firings take place in our brain, do not have the status they would have to have in order to play the role of suitable bridge principles.

Acknowledgements

I would like to thank Antonia Barke for translating this paper into English and Beth Wilson for copyediting it. I am also indebted to Andreas Hüttemann who goaded me into reading Broad's *The Mind and Its Place in Nature* even more thoroughly, and to Christian Nimtz, for his useful comments on an earlier draft of this paper.

Notes

[1] The reason seems to be that Broad did not clearly distinguish laws of interaction from laws of composition.

[2] At least this is generally assumed.

References

Beckermann, A. (1992) Supervenience, emergence, and reduction. In A. Beckermann, J. Kim and H. Flohr, eds., *Emergence or Reduction ?*, 94-118. Berlin and New York: Walter de Gruyter

Beckermann, A. (1997) Property physicalism, reduction and realization. In M. Carrier and P. Machamer, eds., *Mindscapes: Philosophy, Science, and the Mind*, 303-321. Konstanz: Universitätsverlag/Pittsburgh: Pittsburgh University Press.

Beckermann, A., Kim, J. and Flohr, H., eds. (1992). *Emergence or Reduction*? – *Essays on the Prospects of Nonreductive Physicalism*. Berlin and New York: Walter de Gruyter.

Broad, C. D. (1925). *The Mind and Its Place In Nature*. London: Routledge and Kegan Paul.

Hempel, C. G. and Oppenheim, P. (1948). Studies in the Logic of Explanation. *Philosophy of
Science* 15: 135-175. Reprinted in C. G. Hempel, *Aspects of Scientific Explanation and Other* 55a
Essays in the Philosophy of Science, 245-290. New York: The Free Press 1965.

Hüttemann, A. and Terzidis, O. (in press). Emergence in physics. *International Studies in the Philosophy of Science.*

Levine, J. (1983). Materialism and Qualia: The explanatory gap. *Pacific Philosophical Quarterly* 64, 354-361.

Levine, J. (1993). On Leaving Out What It's Like. In M. Davies and G.W. Humphreys, eds., *Consciousness: Psychological and Philosophical Essays*, 121-136, Oxford: Basil Blackwell.

McLaughlin, B. (1992). The rise and fall of British emergentism. In Beckermann, Kim and Flohr, eds., *Emergence or Reduction?,* 49-93.

Stephan, A. (1993). C. D. Broads a priori-Argument für die Emergenz phänomenaler Qualitäten. In H. Lenk and H. Poser, eds., *Neue Realitäten—Herausforderungen der Philosophie. Sektionsbeiträge I zum XVI. Deutschen Kongreß für Philosophie*, 176-183. Berlin: Allgemeine Gesellschaft für Philosophie in Deutschland.