

IMPORTANT: When citing this article, please refer to the print-version:

Erkenntnis. (48)1998, 133-151.

GODEHARD BRÜNTRUP

**IS PSYCHO-PHYSICAL EMERGENTISM COMMITTED TO DUALISM?
THE CAUSAL EFFICACY OF EMERGENT MENTAL PROPERTIES**

1. Introduction

Emergentism is a theory of the mind-body relation that was made popular in the early decades of this century by philosophers like Broad, Alexander, and Morgan. (Alexander: 1927; Broad: 1925; Morgan: 1923.) It promised a middle-position between the extremes of reductive mechanistic theories and scientifically questionable vitalistic explanations. Due to the successes of reductive explanations in science, such as quantum mechanical explanation of chemical bonding in terms of electro-magnetic effects and the explanation of genetics in terms of molecular biology, emergentism has lost many of its adherents. Only recently, mostly due to the problems of giving a reductive explanation of mental entities, is emergentism being revitalized.¹

Emergentism argues that mental properties are

- (a) supervenient on physical properties,
- (b) not reducible to physical properties, and
- (c) causally efficacious.

In this respect it does not differ from the claims of non-reductive physicalists. Many emergentists argue, however, that their position is not a materialist position. It is meant to provide a middle road between physicalism and classical dualism, esp. Cartesian substance dualism: the slogan being 'dualism no – mentalism yes'. Emergentism is meant to overcome the shortcomings of physicalism (the inability to allow for genuine mental causation), while avoiding dualism and its notorious problems with psycho-physical interaction. To explicate the reasons behind the renewed interest in emergentism, it will be necessary to state the mind-body-problem in causal terms, and then show why the received view founders on the rock of mental causation. Functionalism will be analyzed as the paradigm case of the received view. It will be shown that functionalist theories leave no causal role for the mental as mental.

Emergentism, in contrast, tries to secure causal efficacy for higher-level properties, including mental properties by setting them apart from mere functional or structural higher-level properties. Critics of emergentism argue that this claim is incompatible with the principle of the causal closure of the physical. Emergentism, therefore, collapses into a form of outright dualism and thereby inherits the dualist's problems with psychophysical interaction: Absolutely novel mental properties, emerging from an entirely physical realm exert downward causal influences on this physical basis. Even though this may not be a full substance dualism in the specific Cartesian sense, an emergentism of this kind fails to provide a metaphysically significant alternative to classical interactionist dualism. Interactionist dualism holds the following claims:

- (1) Everything real is physical or mental.
- (2) The physical and the mental are entirely distinct.
- (3) There is bidirectional causal interaction between the mental and the physical.

Emergentism seems to be committed to all three of these principles. In this paper I wish to analyze whether emergentism is necessarily committed to a dualism of this kind or whether emergentism can make good on its claim to provide an alternative to dualism.

IMPORTANT: When citing this article, please refer to the print-version:

Erkenntnis. (48)1998, 133-151.

1.1 The Mind-Body-Problem

In one of its traditional forms the mind body problem is stated in *causal* terms. For the following argument I assume a metaphysical-realist account of causation. Causation is a relation between events that exist mind-independently and that are causally efficacious in virtue of some of their properties. Moreover, I hold a nomic-subsumptive model of causality, requiring the subsumption of causally related events under natural laws. The following set of three incompatible claims is a classical version of the mind-body problem, stated in causal terms:

- [1] The physical world is causally closed.
- [2] The causal closure of the physical world entails the causal inefficacy of mental entities.
- [3] Mental entities are causally efficacious.

Giving up one of these three principles resolves the contradiction. Recent discussions in the analytic philosophy of mind were focused on the negation of principle [2]. Negation of [1] seems to imply interactionist dualism with its notorious problem of explaining the *modus operandi* of such interaction. Psychophysical interaction endangers the well-established network of basic physical laws. Unpredictable mental interference undermines the generality of physical laws. Crossing chains of mental and physical causation would make prediction of many physical events impossible to even an idealized physics. Negation of principle [3] provokes 'epiphobia' (J. Fodor), the fear of rendering mental phenomena useless and causally inefficacious. Such entities would be hardly more than decorative ornaments of the furniture of the universe. Why would mental properties evolve, if they had no causal role to play? How could they be known if they could not even affect us causally? Only in an abstractionist or eliminativist framework does the negation of [3] seem to make perfect metaphysical sense. Given these alternatives, a plethora of theories have been developed which claim the compatibility of the causal closure of the physical realm and the causal efficacy of the mental, thereby giving up the, *prima facie*, intuitive principle [2].

1.2 A Metaphysical Image

The metaphysical image underlying most of these views is a layered model of reality that views the world as structured into different levels in a hierarchical order. The bottom level consists of elementary particles, the basic constituents of matter. As we go up the hierarchical structure, we move from atoms to molecules, cells, ensembles of cells, primitive organisms, ultimately to complex organisms and social groups of such organisms. Each level has properties that are characteristic for the particular entities of that level. Consciousness, for example, seems to be characteristic of complex organisms. The model rests on a mereological thesis: a macro-level property M of an object O is constituted by micro-level properties of parts of O which do not have M. Since there are no non-physical particulars according to this picture, it may be considered physicalist in a broad sense of the term. Entities that belong to a certain higher level can be completely decomposed into basic physical parts.

The salient metaphysical problem with this image is the exact ontological status of higher-level properties, esp. mental properties. Given the physicalist framework, the higher-level macro-properties of an entity cannot be something over and above the elementary physical parts of which it is composed. The macro-properties are constituted by the micro-properties. Two entities indiscernible at the micro-level are indiscernible at the macro-level. Micro-properties determine completely the macro-properties (micro-determination). Macro-properties, on the other hand, cannot causally determine the behavior of micro-level events without interfering with the basic causal laws at the micro-level. Mental macro-properties, for example, cannot causally determine

IMPORTANT: When citing this article, please refer to the print-version:

Erkenntnis. (48)1998, 133-151.

micro-level events without breaching the causal closure of the physical. In a given macro-object, the macro-properties cannot add any causal potentialities exceeding those of the various micro-properties. If mental properties are macro-properties in this sense, they are causally inefficacious qua mental properties.

A problem is already foreshadowed here: mental macro-properties are usually identified by their *causal-functional* role. They are functional properties. Because many macro-properties specify structures with multiple realization (like having a desk-like shape), in the debate on emergentism they are often referred to as structural properties. Yet structural properties and functional properties are not identical. It is a structural property of a desk to be supported by legs, and it is a functional property of a desk that it is being used as a means to support one's body in the process of writing. Is 'being a desk' a functional or structural property? It seems reasonable to conclude that many higher-level properties are a combination of structural and functional aspects. In what follows I will focus more on the functional aspect of higher-level properties, since mental properties are most commonly construed as functional properties. For systematic reasons I will be using functionalism as a backdrop for the discussion of emergent properties. But what is going to be said about the causal efficacy of higher-level functional properties can with little modification be applied to higher-level structural properties as well. This is important because emergent properties are often construed as structural properties. For now, however, I will focus on functional properties.

A typical functional property would be the property of being an airfoil. An airfoil is an aircraft part or surface, such as a wing, propeller blade, or rudder, the shape and orientation of which controls stability, direction, lift, thrust or propulsion. This property is clearly identified by its causal role, it seems to be causally efficacious, and it is a higher-level property of a particular that allows for multiple realization and that is constituted by other properties of proper parts of that particular. Within the given space limitation, it is impossible to do justice to all the different theories that have been put forward to account for the causal efficacy of functional macro-properties in a physicalist framework to make intelligible the claim that the mental has a causal role to play in a causally closed physical world. I wish to argue that the debate has demonstrated that only a reductionist form of physicalism can safeguard at least some very limited causal efficacy of mental properties. In several papers on non-reductive physicalism Jaegwon Kim has shown that the concept of non-reductive physicalism leaves no room for mental causation. (Reprinted in Kim: 1993.) Arguing somewhat along those lines, it can be shown what exactly sets emergentism apart from physicalist theories denying principle [2].

2. Two Kinds of Functionalism

The most influential theory that negates principle [2] is, arguably, functionalism in all its varieties. It has been debated whether functionalism is a reductive or a non-reductive strategy. The reason for this ambiguity is that functionalism comes in two flavors. The crucial difference between the two is the status of the causal functional role of a state type or property. According to the first, the causal role constitutes the essence of the state types or properties in question. According to the second flavor, the causal role is only used to identify certain state types or properties, it does not provide their essence. In the first case the causal-functional role is a criterion of identity, in the second a criterion of identification. I shall label the first form of functionalism 'strong functionalism without psychophysical reduction'. The second form, accordingly, will be called 'weak functionalism with psycho-physical reduction'.

IMPORTANT: When citing this article, please refer to the print-version:

Erkenntnis. (48)1998, 133-151.

2.1 Strong Functionalism

Strong functionalism claims that there is a relation of type-identity between mental states and functional states. Mental states are second-order properties specified in terms of functional roles of first-order properties. For example: The property of having the belief that it is raining. This property is identical with the property of having some first order property P, such that being causally affected in a specific way (raindrops, etc.) causes one to have P, and having P causes one to act in a specific way (like opening an umbrella). More generally speaking, this functionalization of a second-order property is achieved by formulating a Ramsey sentence which quantifies into a conjunct of sentences in which the causal connections of some variable x are expressed. This procedure abstracts from the individual occupants of the causal role. The mental predicate is identified with whatever the Ramsey sentence claims to exist. According to the strong functionalist view, mental properties are causal roles, defined in terms of causal relations holding among first-order physical properties. A welcome consequence of this account is the multiple realizability of functional states. Mental states may be reduced to functional states (ideally by theoretical identification). They cannot be reduced to physical states. At best each particular mental event token might be token-identical to some physical event token. But even that is an additional constraint not strictly implied by the theory. Even if in our world w* all functional states are realized in physical states, there are possible worlds in which this is not the case. In a neighboring possible world, a functional state F might be realized by ordered pairs of states, where each pair consists of a non-physical and a physical state. There is even a possible world in which a given functional state F is realizable by a number of non-physical states. The ultimate constituents do not matter. Only the higher level structure matters. As Putnam (in his functionalist period) put it: "The same explanation will go in any world (whatever the micro-structure) in which those higher level structural features are present. In that sense this explanation is autonomous The conclusion I want to draw from this is that we have the kind of autonomy that we are looking for in the mental realm And we need no mysteries, no ghostly agents, no *élan vital* to have it." (Putnam: 1975,296.)

Putnam seems to argue that higher level functional structures are ontologically independent of their realization base. With a closer look it becomes obvious that he is only talking about explanations. He is interested in an autonomous epistemic level. To interpret that autonomy as causal independence means smoothing over ontological cracks. If functional properties can be realized in multiple ways, they cannot be causally efficacious, as the following argument shows:

- (1) Functionalist thinking is the process of identifying one property that covers many realizations.
- (2) If one property covers many realizations, it is they and not it that are causally efficacious.
- (3) Functionalist thinking is not the process of identifying properties that are causally efficacious.

The airfoil is a clear example of that fact. There may be non-basic *ceteris-paribus* laws governing the behavior of airfoils. But non-basic laws rely on mediating mechanisms which they do not articulate. It is here (the underlying mechanism) where the causal work is being done. The functional property is a mere resultant macro-property. It cannot introduce a pattern of behavior at the micro-level differing in kind from the sort that would occur in its absence. Another line of thought supports this argument: The higher-level property is specified by a causal role. The causal role itself cannot be a cause. The cause is first-order, the causal role is second-order. They cannot be the same thing. So-called 'higher-level causal powers' should be understood as just more general and abstract characterizations of the underlying fine-grained causal powers of that particular realization of the macro-property. If this analysis is correct, strong functionalism cannot account for

IMPORTANT: When citing this article, please refer to the print-version:

Erkenntnis. (48)1998, 133-151.

the causal efficacy of functional (esp. mental) properties and states. All it can do is save the autonomy of an explanatory level (it preserves a way of talking we want). Explanations are better if they are more general. It is part of the pragmatics of explanations that abstracting from a specific case to general patterns is a useful strategy.

Against its own original intent, strong functionalism, therefore, resembles the abstractionist-instrumentalist view of mental properties advocated by Daniel Dennett. Mental entities are useful abstractions from the underlying physical micro-structure, but do not exist in a strict metaphysical-realist sense. It is exactly this insight that gave rise to the second form of functionalism.

2.2 Weak Functionalism

Weak functionalism differs from strong functionalism in both structure and objective. It advocates outright psycho-physical reduction. The causal role is used to bridge the gap between the mental and the physical so as to allow for identification of properties (type-identity, or at least something very close to that). This theory (as proposed by David Lewis and more recently Jaegwon Kim)ⁱⁱ is essentially the old central state identity theory refined by a functionalist ingredient. Characterizing a mental concept like pain by its functional-causal role serves the purpose of finding a first-order physical property that occupies the same causal role. The basic line of reasoning goes:

Mental state M = occupant of causal role CR

Physical state P = occupant of causal role CR

M=P.

The mental state is not identical with a causal role, it is only being identified (epistemically) by its causal role. The mental state is not a functional state, it is a physical state of the organism. Functionalization is here an epistemic-conceptual tool for *a posteriori* identifications. It is explicitly denied that the functional level has any ontological autonomy. The functional level is merely a conceptual level. Second order functional concepts refer to sets of first-order physical properties conjunctively. The functional property is not, strictly speaking, a property in its own right, not even a conjunctive property. Quantifying over properties does not create new properties. Functional properties, like the text book example 'dormitivity', therefore, do not exist mind-independently in nature. What really exists are the realizations of those functional properties, for example different chemicals that induce sleep in humans. Even familiar mental properties or states, such as pain, do not exist as such, only species-specifically as pain in humans, pain in mice, pain in aliens from Alpha Centauri. We may use the general functional terms successfully for pragmatic reasons. Theoretical identification, however, is only possible with more specific terms: M-in-A = P (A being a species). It is only this epistemically justified functional level that allows for multiple realization. The mental state is not realizable in multiple ways.

Reduction is achieved by functionalizing an intrinsic property like pain, and then empirically finding the occupant of that causal role relative to some species. That way we can achieve local species-specific reduction. This account of reduction differs from Ernest Nagel's account which has become more or less the text book theory of reduction. Nagel's model stresses the derivation of laws of one theory from another theory, using 'bridge laws' that connect the predicates of the theories. According to Kim, however, it is not necessary for successful reduction to find property connecting bridge laws. In order to successfully functionalize the property in question, it is sufficient to specify its causal-nomic connections, and then find a micro-property that satisfies the specified causal role. If this strategy works for psycho-physical reduction, the causal role of the mental is preserved. Here it is precisely the fact that mental properties are reducible to physical properties, that guarantees their causal efficacy. Of course, there is no causal autonomy of the mental macro-properties in this

IMPORTANT: When citing this article, please refer to the print-version:

Erkenntnis. (48)1998, 133-151.

reductionist model. Accordingly there is no problem of causal exclusion. The causal closure of the physical remains untouched. The key question is, of course, whether mental properties can be functionalized appropriately to be reduced in this way.

According to the received view there are two basic concepts of mind: the relational (or psychological) and the phenomenal. The relational concept portrays the mind as the causal basis of behavior. A mental state is characterized by what causal role it plays in the behavior of the organism. What counts here are the extrinsic-relational features of the mental – its information processing capability. The phenomenal concept of mind, on the other hand, characterizes a mental state by what it is like to be in that state, how it feels to be in that state. This is the concept of a mental state as a consciously experienced phenomenal quality. Here it is the intrinsic property that counts. If the mind is characterized by the relational concept only, it may well be that all mental states and properties are, in principle, functionizable in the appropriate way to allow for reduction. Of course, there has been widespread discussion of exactly how much of the wide variety of mental phenomena can be covered by the relational concept of mind. Searle and Putnam have forcefully argued that functionalist reductions of intentionality have failed. Even if this is still an open question, many would agree that there are at least some mental phenomena that escape the relational approach to the mental, esp. qualia. If that is the case, the reductionist program fails as an all-encompassing strategy. The only remaining option may be to eliminate qualia, setting up a Procrustean device where everything intrinsic is cut off. Kim has argued that this strategy reminded him of the so-called 'Vietnam Metaphor': Saving a village by destroying it. Mental causation may be saved, but mentality (or at least part of it) is lost in the process. (Kim: 1993,367.)

3. Emergentism

The renewed interest in emergentism is, at least to some extent, the attempt to overcome the limitations of a functionalist approach to higher-level properties by securing a genuine causal role for the mental as emergent phenomenon. The main intention behind emergentism is the idea that there must be a conceptual middle-road between the extremes of physicalism and Cartesian substance dualism. One of the key questions is whether this middle ground can still be subsumed under the negation of principle [2]. I shall argue that it cannot be subsumed under principle [2], because emergentism cannot hold on to the causal closure of the physical realm. According to what has been said above, functional properties as such cannot play a decisive role in this new conceptual framework of emergentism. They lack ontological significance qua functional properties and are of mere epistemic value. Their ontological status can only be secured by restricted reduction to physical properties. Emergentism, however, is committed to non-reductionism. Emergent properties must be intrinsically different from functional properties.

In an essay on emergent properties O'Connor gives a comprehensive definition of property emergence:

“Property P is an emergent property of a (mereologically-complex) object O iff:

- (1) P supervenes on properties of the parts of O;
- (2) P is not had by any of the object's parts;
- (3) P is distinct from any structural property of O; and
- (4) P has direct ('downward') determinative influence on the pattern of behavior involving O's parts.” (O'Connor: 1994, 97f.)

Supervenience, novelty, non-structurality, and causal influence are the key elements of emergence. The somewhat unusual term 'non-structurality' needs some clarification. An example of a structural property according to O'Connor would be an object's chair-like shape. It is a mere resultant of the

IMPORTANT: When citing this article, please refer to the print-version:

Erkenntnis. (48)1998, 133-151.

underlying micro-structure. Its causal potentialities are completely reducible to the causal potentialities of its realizer. Structural properties are in that respect exactly like functional properties. What has been said about functional properties applies to them as well. Within the line of the argument given above, O'Connor's account ought to be extended to include functional properties. The relevance of this extension is particularly salient for the emergence of mental properties, because according to the standard account mental properties are considered to be functional properties. With this addition O'Connor's definition features all the metaphysically relevant aspects of property emergence. On the basis of the analysis of functional properties given above, it is important to emphasize that emergent properties, according to the extended definition, are intrinsically different from mere functional properties. But what is it that sets them apart from other higher-level properties?

3.1 Two Notions of Supervenience

One possible way to set emergent properties apart from other macro-properties is by looking at their respective relations to the micro-level properties. Emergent properties supervene on physical subvenient properties, but so do functional and structural properties. A further distinction is required. Supervenience as a metaphysically meaningful relation is construed to imply dependence and determination. Supervenient macro-properties are dependent on and determined by their subvenient micro-properties. Any object x in any possible world and any object y in any possible world that have the same low-level properties also have the same functional and structural properties. The supervenience relation is strong:

For any worlds W_j and W_k and for any objects x and y : if x has in W_j the same subvenient properties that y has in W_k then x has in W_j the same supervenient properties that y has in W_k .

This does not entail the rather strong claim that all natural laws at the macro-level are entailed by the laws at the most basic level. What is being claimed here is weaker. All macro-level facts, including functional and structural facts, are entailed by all the micro-physical facts. A Laplacean super-intelligence could figure out all the functional and structural macro-properties, once presented with all the micro-physical facts. If P is a macro-level property of some object O , then P supervenes with logical necessity on O 's micro-physical parts. Even God could not create a world that was indiscernible from ours at the micro-level but featured different structural or functional macro-properties.

Van Cleve has argued that emergent properties differ from mere resultant properties because they are not logically but merely nomologically implied by their subvenient base. (Van Cleve: 1990.) This is an important distinction. The phenomenal aspect of mentality does not seem to be logically supervenient on micro-physical facts. It seems logically conceivable that there exist a world indiscernible from the actual world in its physical properties, while lacking any conscious experience. A zombie which is identical to me molecule for molecule, will certainly be functionally identical to me, as well. If in the actual world the phenomenal character of mind is only nomologically supervenient on the physical, then phenomenal properties are metaphysically distinct from structural supervenient properties like the shape of an object, or even its information processing capabilities.ⁱⁱⁱ Even if all the physical facts have been fixed, the emergence of consciousness is not implied with nomological necessity. A further fact is required to establish this nomological connection. This squares well with the claim that the existence of emergent properties could not be predicted by even a perfect knowledge of the underlying physical facts alone. Emergentism has often been characterized by the claim that emergent properties could not be deduced even from a complete knowledge of their underlying micro-structure. Beckermann

IMPORTANT: When citing this article, please refer to the print-version:

Erkenntnis. (48)1998, 133-151.

suggests a definition of this kind.

For any system S which has a micro-structure consisting of micro-components connected to each other by a certain relation, in short $(C_1, \dots, C_n; R)$: 'F is an emergent property of S iff (a) there is a law to the effect that all systems with this micro-structure have F, but (b) F cannot, even in theory, be deduced from the most complete knowledge of the basic properties of the components C_1, \dots, C_n ;' (Beckermann: 1992,104.)

Emergent properties might thus be distinguished from simple functional or structural properties by the nature of their supervenience relation to the physical realm. The epistemic claim of emergentists, according to which emergent properties are at first (prior to their occurrence) unpredictable, may then be based on an ontological claim about the nature of the supervenience relation between emergent and base properties. O'Connor has maintained that this will cause an inflation of emergent properties, since there are so many macro-properties that supervene with nomological necessity only on the micro-properties. It is hard to see how this claim could be substantiated. It seems that there are just very few properties in the universe that do not supervene on the micro-properties with logical supervenience. Conscious experience may be almost unique in its failure to supervene with strong modal force. Facts about biology, sociology or economics seem to be logically supervenient on the physical micro-properties, unless they somehow implicitly refer to phenomenal facts. One of the greatest challenges for the emergentist will be to substantiate the claim that there is a multitude of causally efficacious macro-properties not logically supervenient on the micro-structure. Phenomenal conscious experience is intrinsic and non-relational. The emergentist will be faced with the tough problem of securing a causal-relational role for it. If it turned out to be the only emergent property in nature, the appeal of emergentism would be seriously lessened.

It is a tricky question whether facts about causality at the macro-level are logically supervenient on the micro-level. A Humean might argue that even a Laplacean demon could not infer facts about causation if presented only with the set of all physical facts. An emergentist will argue that not even complete knowledge of all physical facts at the micro-level will be sufficient to infer facts about causation at the macro-level.

But leaving these worries about the extent of logically supervenient properties aside, it is obvious that, if the emergentist were to include logically supervenient properties into the set of emergent properties, a serious problem of reducibility would arise. Logically supervenient properties are, at least in principle, reducible to their subvenient base properties. A human observer may not be able to successfully reduce the higher-level properties of economics to lower-level physical properties, but a being with enormous cognitive powers will be able to perform the reduction, if facts about economics are really logically supervenient on facts about physics. It is therefore essential for the emergentist to argue that emergent macro-properties are not logically supervenient on their subvenient micro-properties. Otherwise emergentism could not hold on to non-reductivism. Whether this constraint limits emergent properties to phenomenal properties is an open question. The British emergentists clearly thought that there was a wide variety of emergent configurational forces governing the behavior of many higher-level systems, especially live organisms.

3.2 The Causal Efficacy of Emergent Properties

Novel causal influence distinguishes emergent macro-properties even more obviously from mere functional or structural macro-properties than the modal strength of the supervenience relation. The claim that the causal efficacy of an emergent property is irreducible to that of the micro-properties on which it supervenes, has been a centerpiece of all emergentist theories. The mental realism and

IMPORTANT: When citing this article, please refer to the print-version:

Erkenntnis. (48)1998, 133-151.

non-reductionism implied in emergentism leave epiphenomenalism as the only viable alternative. But for the reasons mentioned above, this alternative has little attraction. Emergentism is committed to the causal efficacy of emergent properties, even to downward causal powers. Emergent properties may be causally efficacious only within their respective emergent levels of the layered ontology: that is, emergent entities may be causally necessitated only directly by other emergent entities at the same ontological level. But this contradicts the claim that each emergent property is causally necessitated by its physical realizer. Emergentism, then, is committed to downward causation.^{iv} An emergent property is causally efficacious by causally interacting with micro-level entities. It thereby indirectly causes the emergence of other high-level properties by causing their low-level realization base. If emergent properties had this kind of causal efficacy, they would be precisely distinguishable from other supervenient properties, such as mere functional properties.

The idea of new causal forces at the macro-level is metaphysically problematic. But this assumption is the centerpiece of emergentism and there is no way to construe a metaphysically significant version of emergentism without it. But the claim may be softened by assuming that the macro-level emergent causal forces are not completely novel. If proto-forms of these forces are already present at the lower levels the mereological problem disappears: how can the whole have properties that are not had, even to a lesser degree, by any of its parts. The relevance of this move becomes especially obvious in the case of emergent mental properties. Causal autonomy of the mental is incompatible with the causal closure of the physical. Downward mental to physical causation implies the negation of principle [1]. Emergentism, it seems, is a dualist theory. Although it claims, in opposition to full substance dualism, the ontological dependency of mental entities on physical entities, it does not escape the main problem of dualism: explaining the *modus operandi* of the causal interaction of two completely different sorts of entities. In the case of emergentism the question of how a genuinely novel property could arise from a basis that features no proto-form of this property is especially pressing. This idea of emergence should be treated with extreme caution. Emergentism, however, is a conceptually coherent position. It certainly is a logical possibility. But the emergentist does not argue on mere *a priori* grounds. The problem, therefore, lies exactly in the empirical claim, that there are novel causal powers. To introduce such powers seems to be an ad hoc assumption unnecessarily inflating the ontology with new causally efficacious entities.

The opponents of emergentism claim that there is currently not the least amount of empirical evidence that there are higher-level configurational forces at work in the phenomenon of life or the phenomena of consciousness. (McLaughlin: 1992,90f.) The success of reductive scientific strategies in chemistry and biology makes it highly unlikely that there are novel forces in nature which emerge only in the, from a cosmic point of view, rare case of complex organisms. Brian McLaughlin argues:

“As truly remarkable as it is, it seems to be a fact about our world that the fundamental forces which influence acceleration (the electro-magnetic weak force and the strong force) are all exerted at the subatomic level.” (McLaughlin: 1992,91.) O'Connor counters that this tough-minded bold claim is sheer bluff. (O'Connor: 1994,99.) But all O'Connor has to offer in opposition is the negative claim that, while there are no widely-accepted theories committed to the existence of such properties, contemporary scientific knowledge is insufficient to conclusively rule out the existence of emergent properties. If emergentists cannot provide better evidence for such existence, then hardly anyone will be willing to pay the price of introducing new downward causal forces based on a mere promise of future scientific progress. After all, emergentism is arguing against physicalism. At least some of the downward causation is direct mental to physical causation, breaching the causal closure of the physical. In the absence of direct empirical evidence for emergent mental properties, it is precisely this widely-accepted causal closure principle that the emergentist must undermine to give

IMPORTANT: When citing this article, please refer to the print-version:

Erkenntnis. (48)1998, 133-151.

his or her position at least some indirect plausibility.

4. Dual Aspect Theory

I wish to argue that this is best achieved by giving up the constraint that an emergent property of an object may not also be had by any of the object's parts. This constraint ought to be weakened as follows: An emergent property may not be had to the same degree by any of the object's parts. If the emergent property could be had, to a lesser degree, by some of the object's parts, the dualist dichotomies in emergentism will disappear. The difficult notion of downward causation will remain, but it will not be burdened with the additional problem of the causal closure of the physical that the interactionist-dualist has to deal with. In the case of mental emergent properties, it is the idea of proto-mental properties that undermines the principle of causal closure of the physical, since it undermines the idea of an exclusively physical realm. But even such an indirect argument for the possibility of emergent properties is hard to maintain without producing at least some weak empirical evidence, which would give us a rough idea of how this claim might be corroborated in the context of modern science. Could there be any configurational forces capable of downward causation in nature? And more specifically: How could one substantiate mind-like forces of this sort?

4.1 Configurational Forces

If no such empirical evidence can be produced, one feels compelled to accept Brian McLaughlin's verdict that emergentism went wrong for 'deep empirical reasons'. (McLaughlin: 1992,91.) He acknowledges, however, that downward causation might be a nomological possibility. (McLaughlin: 1992,53.) He even gives some examples for what may count as downward causation by a configurational force in modern physics: Bohm's quantum potential and Einstein's field equations of general relativity. Bohm's case is a very interesting one, and a closer look will reveal that his idea of the quantum potential does not only introduce a new configurational force capable of downward causation, it also questions the very idea of a causally closed physical realm. With all due caution, Bohm's theory may serve as an illuminating example of a strategy that has the potential to overcome the dualist dichotomies threatening emergentism. It is important to note that nothing hinges on the choice of this particular example for the conceptual coherence of the argument presented here. It is tempting to use this illustration because even such an outspoken critic of emergentism as Brian McLaughlin concedes that this may be a nomologically possible case of downward causation by a configurational force. But the example is only that: a remote empirical possibility, explored in what follows solely for its informative philosophical implications.

Although I cannot go into the details of Bohm's theory, I believe, a few reminders may highlight the essential features: In the classic two-slit experiment, electrons passing through a screen with two adjacent slits produce an interference pattern on the other side, which looks very much like waves interfering with each other. The standard interpretation of quantum mechanics is unable to calculate the individual trajectories of the particles, and can provide only probabilities instead. In contrast, the wave-like quantum potential in Bohm's theory does enable us to calculate the set of individual trajectories producing that characteristic pattern. Consider an electron coming towards the slits, preceded by its quantum wave. The wave passes through both slits and produces an interference pattern. Each electron follows a well-defined path acted upon not only by the classical potential but also by the quantum potential. Thus the quantum potential is a kind of pilot-wave governing the behavior of the particle. Clearly it is a configurational force. It will contain things like the slit width,

IMPORTANT: When citing this article, please refer to the print-version:

Erkenntnis. (48)1998, 133-151.

the distance between the slits, and the momentum of the particle. It conveys information about the environment of the particle. The field from which the quantum potential arises is very different from any other field we know of. Bohm emphasizes that the quantum potential depends only on the form of the wave (= the information encoded in the wave). So that it can be strong even if the wave amplitude is weak. Changing its amplitude does not affect its causal powers. In this it differs from all other field phenomena. The quantum potential arises from a field that is not like an electromagnetic field. Bohm suggests to rather call it an information field: "One may think of the electron as moving under its own energy. The quantum potential then acts to put form into its motion, and this form is related to the form of the wave from which the quantum potential is derived." (Bohm: 1990,279.) Bohm uses the notion of 'active information' for this new causal force. This information at the quantum level is non-local. Even distant parts of the environment can affect the motion of the particle. The phenomenon of non-locality at the quantum level is thereby explained. The information is active only when it enters into the activity of the particle. Bohm draws a comparison with a radio wave, the informational content of which becomes active only through a receiver. The implications for the philosophy of mind are obvious: "Thus, for example, when we read a printed page, we do not assimilate the substance of the paper, but only the forms of the letters, and it is these forms which give rise to an information content in the reader which is manifested actively in his or her subsequent activities. A similar mind-like quality of matter reveals itself strongly at the quantum level, in the sense that the form of the wave function manifests itself in the movements of the particles." (Bohm: 1990, 281.) Paavo Pylkkänen has argued that the key point in Bohm's quantum ontology is that physical effects have a qualitatively new kind of cause: "It is thought that an informational property of the electrons cause them to behave in the way they do This is a new idea in physics, and clearly implies that our concept of causation in the universe has been extended in a significant way." (Pylkkänen: 1992, 123.) We may call this new force 'physical', but then we have tacitly expanded our notion of what is classified as physical. Typically, mental causation is understood as the causal efficacy of an informational content. According to Bohm's view, then, there is a primitive form of mental causation even at the quantum level. "It would thus seem that physics itself is calling for a change in our habitual way of distinguishing physical properties from informational or mental properties in some absolute metaphysical sense." (Pylkkänen: 1992, 123.) The notion of the causal closure of the physical is then undermined because the notion of an exclusively physical realm is undermined. If there can be qualitatively different configurational causal forces at the quantum level, the existence of some form of active information in the micro-structure of the brain should not be excluded on *a priori* grounds.

4.2 A Similarity Relation

So much about Bohm. His theory is lacking general acceptance and, as an empirical theory, cannot serve as the foundation of a general ontological argument. Nevertheless, this example illustrates nicely how the main objective for an emergentism, which wishes to avoid classical dualism, ought to be to question the idea of the causal closure of the physical. Such questioning should develop a notion of matter that does not allow for an absolute mind-matter dichotomy. If the very notion of the causal closure of the physical has thus become undermined, the idea of causally efficacious higher level mental properties is metaphysically much less problematic. What the emergentist has to abandon is the idea that mental properties arise from a non-mental and exclusively physical realm. It is precisely this notion of absolute novelty that gives rise to the problem typical of classical dualism: The inexplicable *modus operandi* of interaction between two radically different substances. Moreover, it poses the mereological problem of how a complex

IMPORTANT: When citing this article, please refer to the print-version:

Erkenntnis. (48)1998, 133-151.

mental property could arise without ultimately being based on less and less complex mental properties. Emergentism should therefore abandon the constraint that an emergent property of an object may not be had, even to a lesser degree, by any of the object's parts. In the case of emergent mental properties this leads directly to a form of dual-aspect theory, in which the notion of a causally closed physical world literally makes no sense because even low-level events have proto-mental properties.

It is the task of the philosopher to make conceptual sense of a dual-aspect conception of reality. Since 'dual-aspect' is only a metaphor, this is a formidable task. Even though he himself favors aversion of a dual-aspect theory, Thomas Nagel has conceded that dual-aspect theories possess the slightly sickening odor of something put together in the metaphysical laboratory. (Nagel: 1986, 49.) The most pressing problem is the threat of panpsychism. To avoid the unwelcome consequences of outright panpsychism, the proto-mental properties at the base level must be sufficiently dissimilar to full mental properties found in higher-level systems. Dualistic dichotomies, on the other hand, have to be avoided. The similarity relation allows for a middle way between these alternatives.^v In the pyramid of the layered metaphysical picture, a similarity relation must hold between adjacent levels. But since similarity is not a transitive relation there may be no similarity between the properties of the highest and the lowest levels, which would allow for novelty. This notion of novelty does not imply dualist dichotomies. Because all adjacent levels are connected by the similarity relation, there is no absolute novelty (no gaps). A metaphysical theory of this sort provides a conceptually coherent middle road between Cartesian dualism and materialist monism. Emergentism, then, is not committed to dualism. But there is a price to be paid for providing a metaphysically significant alternative to dualism. It is the acceptance of a dual-aspect ontology, which obscures the clear distinction between the mental and the physical. In the absence of convincing empirical evidence for such a world-view (except for highly disputed interpretations of quantum mechanics and some other problematic cases) the situation for the emergentist is not very promising. But my concern was whether emergentism remains a substantive and original position in the philosophy of mind. It clearly does.

References

- Alexander, S.: 1927, *Space, Time and Deity*, 2 Vols. Macmillan, London.
- Beckermann, A.: 1992, 'Supervenience, Emergence, and Reduction', in: Beckermann et al. (eds.), pp. 94-118.
- Beckermann, A., Flohr, H. and Kim, J. (eds.): 1992, *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*, de Gruyter, Berlin.
- Blitz, D.: 1992, *Emergent Evolution. Qualitative Novelty and the Levels of Reality*, Kluwer, Dordrecht.
- Bohm, D.: 1990, 'A New Theory of the Relationship of Mind and Matter', in: *Philosophical Psychology* 3, 271-86.
- Broad, C. D.: 1925, *The Mind and its Place in Nature*, Kegan Paul, Trench, Trubner & Co., London.
- Chalmers, D.: 1996, *The Conscious Mind - in Search of a Fundamental Theory*, Oxford University Press, Oxford.
- Hasker, W.: 1982, 'Emergentism', in: *Religious Studies* 18, 473-88.
- Kim, J.: 1993, *Supervenience and Mind. Selected Philosophical Essays*, Cambridge University Press, Cambridge.
- Kim, J.: 1996, *Philosophy of Mind*, Westview Press, Boulder.

IMPORTANT: When citing this article, please refer to the print-version:

***Erkenntnis*. (48)1998, 133-151.**

Lewis, D.: 1983, *Philosophical Papers Vol. 1*, Oxford University Press, Oxford.

Lewis, D.: 1994, 'Reduction of Mind', in: Guttenplan, S. (ed.), *A Companion to the Philosophy of Mind*, Basil Blackwell, Oxford, pp. 412-30.

McLaughlin, B.: 1992, 'The Rise and Fall of British Emergentism', in: Beckermann et al. (eds.), pp. 49-93.

Morgan, C. L.: 1923, *Emergent Evolution*, Williams and Norgate, London.

O'Connor, T.: 1994, 'Emergent Properties', in: *American Philosophical Quarterly* **31**, 91-104.

Putnam, H.: 1975, 'Philosophy and Our Mental Life', in: *Mind, Language, and Reality ; Philosophical Papers Vol. 2*, Harvard University Press, Cambridge, MA., pp. 291-303.

Pylkkänen, P.: 1992, 'Mind , Matter, and Active Information. The Relevance of David Bohm's Interpretation of Quantum Theory to Cognitive Science' , in: *Reports from the Department of Philosophy of the University of Helsinki 2*, University of Helsinki, Helsinki.

Nagel, Th.: 1986, *The View from Nowhere*, Oxford University Press, Oxford.

Spencer-Smith, R.: 1995, 'Reductionism and Emergent Properties', in: *Proceedings of the Aristotelian Society* **XCIV**, 113-29.

Van Cleve, J.: 1990, 'Mind-Dust or Magic? Panpsychism Versus Emergence', in: *Philosophical Perspectives* **4**, Action Theory and Philosophy of Mind, Ridgeview, Atascadero, pp. 216-26.

Hochschule für Philosophie München
Kaulbachstr. 31a
80539 München
Germany
Godehard.Bruntrup@hfph.mwn.de

- i For recent discussions of the concept of emergence see, for example Hasker: 1982; van Cleve: 1990; Beckermann et al. (eds.): 1992; Blitz: 1992; O'Connor: 1994; Spencer-Smith: 1995.
- ii Lewis: 1983, Part 2; Lewis: 1994; Kim: 1992: 'Multiple Realization and the Metaphysics of Reduction', reprinted in Kim: 1993,309-35; Kim: 1996, chap. 5.
- iii As recently argued in Chalmers: 1996.
- iv As J. Kim has forcefully argued in his 'The Nonreductivist's Troubles with Mental Causation', in Kim: 1993,336-57.
- v For discussion of this issue I am grateful to Gary Rosenkrantz.