Emergent Panpsychism

Godehard Bruentrup

Outline

"Se non è vero, è molto ben trovato—If it is not true, it is well conceived." This quote is usually attributed to Giordano Bruno (1548–1600). Bruno was one of the great panpsychists of the Renaissance. He believed that all matter is permeated by spirit. Panpsychism has always been a player in Western metaphysics, albeit not as prominent as dualism or materialism. The mind-body problem might be inscrutable to us and panpsychism might not be true, but it is a coherent position that deserves more attention in the contemporary debate. In this chapter three claims will be defended:

1. Panpsychism is a genuine and conceptually clearly delineated position in the philosophy of mind. It should be taken seriously as an alternative to the dominant and somewhat simplistic view that the only viable options are physicalism or dualism. It is distinct from idealism and neutral monism as well.
2. Constitutive panpsychism is currently the preferred variant of panpsy- chism. Retaining the causal closure of the physical, this compositional form of panpsychism seems to provide a metaphysics of mental causation that is acceptable within a broadly physicalist framework. It is doubtful, however, whether it can preserve a robust metaphysical account of agency. This is the reason why alternative versions of panpsychism deserve attention.
3. Nonconstitutive panpsychism, a hitherto less developed account, has the resources to overcome some of the problems facing constitutive panpsychism. Even if incorporating the notion of downward causation, it is nevertheless clearly distinct from substance dualism.

Panpsychism as a Distinct Position in the Philosophy of Mind

This section will sketch the most general metaphysical claims of panpsychists and attempt to differentiate panpsychism from the more general intuitions of Russellian monism.

Surprisingly, panpsychism is often accused of being physicalism or dualism in disguise, that is, of collapsing into one of these two positions in the final analysis. The surprising fact that both of these claims are actually advanced can be explained more easily by recalling that two distinct versions of panpsychism are distinguished. One, constitutive panpsychism, is closer to classical physicalism; the other, nonconstitutive panpsychism, is closer to classical dualism. Both versions of panpsychism share basic assumptions that set them apart from these other accounts. Constitutive panpsychism claims that macroexperience is constituted by, grounded in, or realized by microexperience. Nonconstitutive panpsychism claims that unified macroexperience is an emergent phenomenon that cannot be fully captured by the metaphysical concepts of constitution, composition, grounding, or realization.

Ontological Simplicity

Like other metaphysicians, panpsychists are often motivated by certain general and overarching metaphysical ideals of simplicity and austerity: the homogeneity of nature, monism, and a single category ontology. Many physicalists and idealists share the same goal of theoretical simplicity, including proponents of a physicalist event ontology or an idealist Leibnizian monadology. From the point of the panpsychist, physicalism and idealism do, however, pay a substantial metaphysical price by either reducing the mental to the physical or by deflating the physical to the status of a well-founded phenomenon. Taking both the mental and the physical seriously, while avoiding the complexities of substance dualism, is certainly a major motivation for panpsychists. If panpsychism cannot preserve the elegant simplicity of physicalism or idealism, then dualism might well be the more attractive alternative, since it is capable of preserving our common-sense worldview more effectively than panpsychism.

The homogeneity thesis claims that nature is built up from the same kind of basic entities through space, and even through time. Beings in very distant galaxies will be made up of basic elements that are prevalent in our part of the universe. So it will not be the case that some cosmic locations will contain simple Cartesian souls, whereas others contain only elementary physical particles and compositions of particles.

According to the panpsychist, it is also not the case that some cosmic locations contain entities endowed with experience and that the entities in others are completely void of even the simplest form of experience. Some experiential aspect (which might be very primitive and not necessarily conscious) is fundamental to, and ubiquitous in, concrete particulars. Experience (in the widest possible sense) exists not only at some rare level of complexity, such as animal brains, but is in some form present at all levels of nature. The diachronic ideal of homogeneity states the same idea in temporal terms. In the history of the universe there is no clear cut-off point at which experience emerges out of a past that is absolutely void of any experience. Rather, experience was somehow "present at the very origin of things" (James 1983, 152); it is a fundamental feature of nature.

The theoretical intuition guiding these thoughts is the so-called genetic argument. *Ex nihilo nihil fit*, nothing can emerge out of nothing. More precisely, it states that nothing can give something which it does not possess. For something to emerge out of something else, the former must have been present in some diminished form right from the beginning. For example: Even if the solidity and impenetrability of macroobjects cannot be found in the microworld, there have to be physical properties that serve as an emergence base for these macroproperties; otherwise the emergence becomes a mere brute and unintelligible fact. Imagine a Platonic world inhabited only by abstract entities. The idea that in this world some kind of configuration of abstract objects can cause the emergence of a concrete material object seems unintelligible. If there are no spatio-temporal entities in a given world from the outset, no configuration or structure in that world will bring about spatio-temporal particulars. Similarly, the panpsychist claims that the emergence of phenomenal minds out of a world which lacks a scintilla, even the faintest form, of experience, is simply unintelligible. One might also call this the denial of inter-attribute emergence. Say, a given ontology contains these four basic attributes: concrete, abstract, physical, mental. If, for example, a given possible world in this ontology contains at its foundation concrete physical objects only, then neither mental nor abstract entities can emerge in it, at least it cannot be made intelligible how they could emerge. But higher levels of concrete physical objects can emerge in it because this would only amount to a case of intra-attribute emergence.

Monism

For the same reason, panpsychism is often labeled as a kind of monism. But this description might be too coarse-grained. Of course, seen as an alternative to Cartesian dualism, panpsychism is monistic in spirit. There is a possible reading of panpsychism where all physical properties are grounded in mental properties.

If grounding means that the intrinsic properties are the constitution base of the relational extrinsic properties, then we have a metaphysical system in which certain fundamental entities with absolutely intrinsic properties constitute all the remainder of reality — a view that Leibniz famously argued for in his Monadology. The whole of physical space was constructed by giving each monad a spatial viewpoint from which space was constituted as *experienced space*. But then panpsychism simply collapses into idealistic monism. The version of monism that comes closest to panpsychism is neutral monism, but neutral monists take great care to distinguish their position from panpsychism. Neutral monism is often somewhat Kantian in spirit, claiming that we do not know the ultimate nature of the universe but only the mental and the physical arising from that ultimate nature. The neutral is simply not known to us; it is a postulate of reason. This Kantian ignorance about the ultimate nature of things was certainly a dominant motive for Russell's neutral monism. While Russell flirted with the idea that the intrinsic properties of the physical might be mental in nature, his predominant attitude about the intrinsic nature of things seems to have been agnosticism. If panpsychism is labeled as 'Russellian monism,' then — while not being a complete misnomer — relevant differences will be obscured. The panpsychist would argue that an entirely neutral basis that does not even contain some primitive or 'proto' form of mentality would not help us at all in explaining the emergence of the mind. Thus panpsychistic monism is a 'dual-aspect monism,' claiming that there is only one kind of thing but it features physical and mental properties. Dual-aspect monism is not a contradiction in terms, but the concept certainly makes it clear that panpsychism is neither monism nor dualism *simpliciter*. Panpsychism is more complicated than these crude alternatives suggest. In the contemporary debate about panpsychism, the mental part is usually constituted by the intrinsic properties of things (the quiddities), whereas the physical part is constituted by the relational-structural properties. Whether this is still a form of monism in the full sense is disputable. One could also, with some justification, call it 'dualism all the way down.' Panpsychism is not a monism *tout court* as is physicalism or idealism. Even so, it may correctly be labeled 'monist' in the sense that it can be construed as a one-category ontology. Whitehead's panexperientialism is a good example for this. For Whitehead there is only one basic category of entities: events with both a physical and mental pole. He calls events 'actual occasions.' They are the most fundamental entities. Everything else is constructed out of them. The question, however, is whether this construction allows for emergent novelty rather than merely a resultant composition of existing entities. Whitehead was a nonconstitutive panpsychist who allowed for the strong emergence of new individuals. These new individuals were strongly emerging actual entities endowed with a new mental unity and subjectivity.

In the most recent debates, panpsychists have often tried to avoid strong emergence. The attraction of constitutive panpsychism is rooted in its ability to preserve a broadly reductionist picture. The microentities determine all facts. But not all facts are physical facts in the narrow sense. There is a categorical basis that carries the relational properties described by physics; ignoring it is the reductive physicalism's mistake. This move enables the panpsychist to escape the zombie argument against physicalism by claiming that a complete copy of the physical world must include the quiddities and not just the relational properties. Such a metaphysically complete copy of the lowest physical level — that is, a copy which is not only structurally isomorphic but copies the intrinsic natures as well — will give rise to consciousness with the strong necessity that the relation of logical supervenience provides. It might seem that constitutive panpsychism is simply physicalism in disguise. But to think this would be a grave misunderstanding, conflating two senses of 'physical.' To use Strawson's terminology, it is important to conceptually distinguish 'physicalism' from 'physicSalism.' The metaphysical work that is being done by intrinsic mental properties, even at the most basic level of nature clearly distinguishes constitutive panpsychism from physicalism.

The Carrier Thesis

The distinction between physicalism and physicSalism points again to the duality inherent in panpsychism, even if the latter is construed as a one-category ontology. The physical structure is metaphysically incomplete, requiring a categorical base. Only the composition of both the structural and the non-structural will give rise to a real concrete entity. This thesis could be called the 'hylomorphic thesis' or better the 'carrier thesis.' As in Aristotle's metaphysics, relational structure can only exist together with something ultimately non-structural which it configures. But the term 'hylomorphism' is strictly associated with the Aristotelian fundamental duality of form (*morphe*) and matter (*hyle*). In the Aristotelian tradition form requires something which it configures (prime matter). Aristotelian prime matter carries the relational structure of the world. Structure alone is never sufficient for the existence of a concrete particular. Panpsychists often argue in a parallel way, but for them the carrier is not prime matter but something mental or analogous to the mental.

The modern post-Cartesian view of matter reduces matter to something that can be fully described by mathematical relations in algebra, analytic geometry, and calculus. It thus focuses on the structural properties of things. Even the Cartesian idea of matter as mere extension is silent on the question of what is being extended. Extension is merely the abstract notion of the repetition of something — a point that was already made by Leibniz (see Brüntrup 2009, 246).

Leibniz argues that the concept of extension cannot help explicate the nature of the substance that is being spread out, and that, on the contrary, substance is ontologically prior to the repetitive multiplicity of extension (see Leibniz G IV, 467). Formally speaking, extension is the repetitive multiplicity of point-like entities. But the nature of those entities is specified in the sciences simply by the relations into which they enter.

A pragmatic way to avoid these deep metaphysical waters would be to resort to epistemic structuralism, which is antirealist in spirit and remains agnostic about any nonstructural properties of the unobservable. Nature, as we know by empirical investigation, is only nature as it presents itself through the mathematical analysis of sense data. The Russell of "The Analysis of Matter" provides a good example of this view (see Russell 1927). Russell argued for an agnosticism concerning the physical world, with the exception of its purely formal and mathematical properties: "it would seem that wherever we infer from perceptions it is only structure that we can validly infer; and structure is what can be expressed by mathematical logic" (Russell 1927, 254). And: "The only legitimate attitude about the physical world seems to be one of complete agnosticism as regards all but its mathematical properties" (Russell 1927, 270). Higher order properties of physical theories can only be expressed in mathematical terms. Epistemological structural realism limits the scope of scientific realism to exactly these properties. But according to the epistemic structural realist there is an objective world out there that contains unobservable objects, of which we can only know the relational properties. Thus we only know the structure, ultimately only the formal structure, of the world. This idea, unsurprisingly, is much older than Russell. Kant argued that things in the phenomenal world are wholly constituted by their relations. He considers it a conceptual truth, however, that things as objects of pure understanding must have intrinsic properties. Even in the 'postmetaphysical' "Critique of Pure Reason" Kant argues that those intrinsic properties must be analogous to the ones presented to me by my inner sense. "They must be something which is either itself a thinking or analogous to thinking" (Kant CPR, B321). This is a metaphysical argument that most panpsychists would gladly embrace (see Brüntrup 2011, 24).

The classic critique of epistemological structural realism is " Newman's Argument" (Newman 1928, 139–140), which was initially directed against Russell. Newman's Agument is best understood as a *reductio* aimed at epistemological structural realism, showing that it ultimately collapses into antirealism. The existence of a structure is trivially true of a set of objects. According to Newman, a statement describing a certain structure with regard to a number of objects is trivial. Why is it trivial to claim that a set of objects has a (or some particular) structure? Because, for Newman, a structure is purely formal and mathematical, and is furthermore independent of the intrinsic qualities of the objects.

If only the structure is known, then beside what is logically deducible from the properties of the structure, the only thing that can be known is the number of its constituting objects. But if all we know about the objects is their cardinality, that is, if we do not know any properties of the objects that ground certain relations and exclude others, then —mathematically — any system of relations over these objects is as good as any other; all of them are instantiated. Relations are simply sets of ordered sequences of entities. Given the entities, all of those ordered sequences will exist, as a matter of pure mathematics. If Newman's argument is correct, then not only do we have no knowledge of the intrinsic properties of things, we do not even know the objective structure of the world in any realist sense of 'objective.' Scientific realism understood as epistemic structural realism collapses into antirealism.

The same point can be made of Putnam's famous model-theoretic argument (see Putnam 1980). As Newman argues, given a number of objects, any relational structure configuring them is already given. If we picture objects as mere nodes in a relational graph and as having no intrinsic nature, then for each structure there are many different relations between the objects that make true the propositions describing that structure. What is the intended model of the structure? Which is the one and only relational structure of the mind-independent world? There are too many ontological interpretations (models) for our theories. Our scientific descriptions of the world are unable to single out the intended model, that is, the real world. Since science deals only with mathematical structure and not with the relations which are determined by the qualitative intrinsic natures of the relata, we can never know the one true story about the world in a metaphysical-realist way. We have too many 'truths.' Panpsychists are not well advised to take this broadly Russellian route, which leads not only to agnosticism about the intrinsic properties of concrete entities but to antirealism in the philosophy of science.

Intrinsic Natures

Russell's views do not fit smoothly into the metaphysical debates within the contemporary analytic philosophy of mind. There is a significant element of idealism in Russell's thought which simply does not square easily with the metaphysical realism prevalent in the current debate. The Russell of "Our Knowledge of the External World" (Russell 1914) is still present in his writings from the 1920s. For the early Russell the world does not cause sense data; rather, sense data construct the physical object. It is not far from this move to the Kantian claim that the intrinsic nature of physical reality is unknown to us. Genuine panpsychism, as understood here, is a robust metaphysical thesis in which scientifically inaccessible intrinsic mental properties (or properties somehow analogous to mental ones) play an important role in grounding the relational structure of the world.

The relational properties of substances must have an underlying foundation in intrinsic properties. Famously, Leibniz claimed in his "Letter to de Volder" that there is no denomination so extrinsic that it does not have something intrinsic as its foundation (Leibniz G I, 240). If this is correct, then we need ultimate intrinsic properties that carry the entire existing net of functional-relational properties.

This intuition resurfaces in contemporary debates. John Haugeland endorses the traditional view that a substance needs certain properties which it maintains regardless of anything else. He considers the ontological status of the pieces in a chess game—say a rook or a pawn—and claims that their very nature is determined entirely by how they move about in the game in relation to other pieces. "No rook is a substance. Nothing about a rook is determinate, not even its 'rookness,' apart from its participation in a chess game" (Haugeland 1993, 63). The formal definition of a type in a chess game is circular. The nature of each type is completely determined by its set of allowable moves within the game as a whole. The chess game as a whole, however, is defined by the interdependent set of types which play functional roles within it. Each part of the game presupposes the existence of the whole game, and the game presupposes the existence of its parts. Why isn't this circularity of chess categories vicious? How can chess games actually and concretely exist? Classic functionalism has an answer to this question, which is quite similar to the one given by Aristotelian hylomorphism: There must be something distinct from the formal structure that actually grounds that structure in concrete reality. In the case of a chess game we have physically distinct objects that serve as stand-ins or realizers of the relevant types, thus allowing for the existence of concrete tokens of those types.

Of course, there is much more to consider here, like the concrete chessboard or the physical position of the players in space. Without such 'carriers' of the formal structure, the game would remain too incomplete and abstract to exist concretely. The panexperientialist Gregg Rosenberg extends this thought to other, more complex, conceptual systems such as those constructed by scientific theories (see Rosenberg 2004, 234). Such systems too are abstract and circularly defined. Consider cellular automata in computer science. These are like giant chess games. Each cell is defined by its role in the entire system, and the entire system is defined by the cells. Cellular automata may exist as computational systems because there is something external to the formal system that realizes or carries them. The physical states of the hardware are the carriers of the cellular automata as functional systems. Biology as an abstract functional system is carried by the mechanics of molecular biochemistry, psychology by the dynamical properties of the neural system, economics by the needs and desires of individuals.

The crucial question, however, is: What carries the most basic physical level? Physics presents us with a world of interdependently defined functional roles. Are there any properties that can give this circularly defined conceptual system a foothold in concrete reality? This is a puzzling question that Rosenberg calls the 'ultimate-carrier problem.' It is very similar to the question Leibniz raised with regard to Descartes's notion of matter.

There are good scientific reasons to assume that nature has a lower size limit (i.e., a Planck-size scale). In order to avoid an infinite regress of ever more finegrained systems, where each lower structure serves as the carrier of the next higher one, a stopper is needed. Only a property that is intrinsic *tout court* and not relative to a system could bring this about and serve as an ultimate carrier. Are there properties that are not intrinsic to any system but at least partly intrinsic to themselves? The only candidates we know of are phenomenal qualities, or something analogous to phenomenal experience. "Analogous" means that these properties might be vastly dissimilar to higher-level phenomenal properties, for similarity is not a transitive relation. Even if mental properties of adjacent layers of nature are similar, the mental properties of nonadjacent layers may be quite dissimilar. But they must have something in common with the phenomenal properties we experience. They cannot be understood in purely relational terms. One cannot understand the nature of these phenomenal qualities by knowledge of their contextual relations alone. Radical intrinsicness is the very nature of phenomenal qualia. Whatever grounds the structural-relational properties of the world must have this radical intrinsicness. It might well be that our own consciousness is the closest analogue we have to this underlying reality. This is the so-called argument from intrinsic natures for panpsychism. As we have seen, even Kant agrees with this line of thinking, provided that what we are seeking is a metaphysical truth. In his critical philosophy, however, he opts for an epistemic constraint within the boundaries of possible experience. A similar skepticism is present in Russell's agnosticism about the ultimate intrinsic nature of concrete entities. The panpsychist, however, cannot enjoy the luxury of withholding judgment here. For the panpsychist, at least some of the intrinsic properties of things are somehow experiential, analogous to experience, or proto-experiential, but certainly not simply neutral. The Kantian point that such a theory speculates beyond the realm of possible experience is nevertheless well taken. If there are phenomenal properties in nature outside of our own consciousness, then we will never be able to access them directly. Panpsychism vastly expands the problem of other minds. By the same token, if panpsychism is true, then nature is much more similar to us. The conscious human mind is not an alien subject in a mechanistic material universe of Cartesian extended objects.

If our own conscious experience tells us — if only by analogy — something about the deepest levels of the universe, then the hiatus between mind and world may be less deep than modern philosophy has traditionally assumed. The nature of the thing in itself is not completely hidden from us.

We can conclude from these considerations that panpsychism is indeed a robust metaphysical position that is conceptually different from the neutral monisms inspired by the inscrutability of the ultimate nature of things.

Constitutive Panpsychism, Emergence and Mental Causation

Constitutive panpsychism, sometimes also labeled 'compositional panpsychism,' is probably the most discussed position in the current debate. It claims that macroexperience is constituted by, grounded in, or realized by microexperience. In its most attractive form it claims that there is an a priori entailment from microphenomenal truths to macrophenomenal truths. It seems to hit the sweet spot between physicalism and dualism. Due to its type-A a priori character, this position is a stronger modal thesis than the so-called type-B physicalism. It is endowed with all the explanatory power one could ask for. The only sort of identification which the type-B physicalist allows between physical and mental states or events is a posteriori. There is no a priori entailment of phenomenal concepts in physical concepts. In other words: Zombie worlds are perfectly conceivable but they are metaphysically impossible. Type-B physicalism thus rests on a form of modal dualism. There is a clear distinction between conceptual or logical possibility and real or metaphysical possibility. The realm of what is accessible to rationality by logical and conceptual analysis is disconnected from the realm of being, that is, from the realm of what is metaphysically possible. This gap undermines the entire project of employing conceptual analysis to do metaphysics. Metaphysical reasoning is then bound by what is accessible to us by sense experience. Science, not a priori analysis, discovers the nature of things.

In the case of the mind-body problem this leads to a kind of bruteness and opacity of the identity between the mental and the physical. It is in no way transparent to us how a relational structure described by physics necessitates conscious experience. We can discover these psycho-physical relations, but we have no conceptual insight into the nature of the necessity. It is this lack of analysis that motivates panpsychism. From a panpsychist point of view, a fully transparent analysis of the physical concepts will reveal why the basic physical entities will necessitate higher-level consciousness if they are appropriately arranged to promote a highly integrated flow of information.

A fully transparent analysis of the physical concepts will reveal the true nature of the physical, which includes mental or proto-mental intrinsic properties. For the type-B physicalist the most appropriate reply seems to be to claim that physical concepts are fully transparent but that mental ones are not. Talking about phenomenal consciousness does not reveal the true nature of consciousness. The true nature of, say, pain is not captured by the description of what it is like to be in pain. There are two reasons why this answer fails to satisfy the panpsychist. Firstly, for something to be pain, it has to feel like pain. Trying to define the identity conditions for pain without referring to phenomenal concepts seems like dodging the question about the nature of pain, not answering it. The other reason for doubting the type-B physicalist's strategy is given by the scientific principle according to which we understand things by breaking them down into smaller components. We wish to understand how the phenomenal mind, as we know it from human experience, is constituted by the material described by neurophysiology and physics. Type-B physicalism simply cannot provide a satisfying answer to the constitution question, unless it somehow gets rid of the phenomenal character of experience. The constitutive panpsychist, by contrast, provides an elegant answer to this question. To illustrate this, it might be best to return to the so-called cellular automata mentioned above.

Cellular Automata

Gregg Rosenberg developed an argument against physicalism based on the idea of cellular automata (see Rosenberg 2004, 14–30). It requires fewer presuppositions than the zombie argument (like possible worlds, 2D semantics) but is nevertheless fully adequate to express the trouble with physicalism and pinpoint the solution suggested by the panpsychist. Cellular automata are artificial digital worlds consisting of basic particulars called 'cells' in an abstract space. These cells have relational properties connecting them to other cells. Computer modelers define various worlds by giving the cells different properties and then studying their dynamics through consecutive computational steps. This is usually done by defining rules that determine which properties a cell will have at a given time as a function of which properties the neighboring cells had at an immediately preceding time. In simple versions of cellular automata the basic particulars only have simple properties like 'on' and 'off.'

One can build on these humble beginnings and construct more complicated cellular automata that may ultimately mimic physical properties like spin, charge, or mass. The fascinating fact is that despite its rather simple physics, the cellular automaton is enormously versatile, in fact a universal Turing machine. Individual cells join together very quickly to build ever more complex structures and patterns that are sustained over many steps of computation.

The machine seemingly produces endurants (i.e., stable relational patterns) that arise out of a sequence of event-like occurrents (i.e., discrete computational states of the system). These patterns become quite sophisticated, featuring a kind of nontrivial self-replication that is functionally similar to certain structures of living beings (i.e., DNA). For this reason cellular automata are sometimes called 'life worlds.' The basic facts of cellular automata—that is, the ways in which properties are distributed over the grid of cells—necessitate all higher-level structural facts about stable emerging patterns. There is no mysterious strong emergence involved, even though the emerging patterns exhibit new properties that cannot be attributed to individual cells. Everything happens within one strictly delineated ontological scheme. Physicalism can be construed as the thesis that our world is an extremely complex cellular automaton. But then the following problem arises:

(1) The fundamental facts of cellular automata are defined entirely by the dynamic relations among the cells.
(2) Facts of phenomenal consciousness are intrinsic qualitative facts, which cannot be entirely defined by the dynamic relations in which they enter.
(3) Facts about dynamic relations do not entail intrinsic qualitative facts about phenomenal experience—neither a priori nor a posteriori.
(4) Thus, the intrinsic qualitative facts about phenomenal experience are not entailed in the facts about cellular automata.

The question is then: If our world is a cellular automaton, how does the phenomenal mind emerge? The answer is straightforward for the constitutive panpsychist: The individual cells have intrinsic natures that are mental or at least analogous to mentality (that is, proto-mental). It is the composition of these intrinsic natures that explains the emergence of phenomenal minds. The emergence is thus neither brute nor inexplicable. The composition of the cells alone accounts for the weak emergence of higher-level structures and higher-level mentality. There is logical synchronic supervenience between the lower and the higher levels. A perfect copy of all of the cells, including their intrinsic natures, will necessitate higher-level structure and higher-level phenomenal properties. The microdeterministic layered ontological framework of physicalism can be fully retained. The physical level determines all the facts, if 'physical' is taken in the broad sense such that quiddities, that is, intrinsic natures, are included. The beauty of constitutive panpsychism lies precisely in its ability to leave the overall framework of traditional physicalism intact. The macrofacts are synchronically microdetermined. Reductive explanations in the sciences are metaphysically vindicated. Constitutive panpsychism just adds nonobservable intrinsic natures to the scientific image.

These natures do much of the metaphysical heavy lifting in the philosophy of mind, without getting in the way elsewhere by interfering with the physical laws governing observable physical processes. For all pragmatic or instrumental scientific purposes it is perfectly acceptable to abstract away from those intrinsic natures. From the point of science the quiddities are mere metaphysical postulates. Thus even if constitutive panpsychism is true, science can work under the presumption of traditional methodological physicalism. Most significantly, constitutive panpsychism can hold on to the causal closure of the physical. David Chalmers sees this as a distinct advantage of constitutive panpsychism over its cousin, nonconstitutive panpsychism. The latter position requires strong emergence and (possibly) downward-causation and seems therefore prima facie incapable of providing a clear theoretical advantage over (emergent) dualism. But how, exactly, does constitutive panpsychism preserve the causal efficacy of the mental, especially in cases where beliefs or desires cause the movement of bodies?

The Causal Efficacy of the Phenomenal Mind

For the constitutive panpsychist, the causal efficacy of phenomenal properties does not rely on their being directly involved in causal relations. Rather, they are thought to be efficacious because they are essential properties of the entities that do enter into causal relations. Without some kind of intrinsic properties, the entire network of causal relations could not exist. This move is somewhat reminiscent of Davidson's theory of mental causation: Mental properties do not enter into the causal laws, but that—according to Davidson—does not render the mental epiphenomenal (see e.g., Davidson1993). It is not the event *qua* mental or *qua* physical properties that is causally efficacious; it is the entire event, the event as such. Changing the mental properties will yield a different event and thus a different causal story. But doesn't this leave the phenomenal properties as being epiphenomenal? The entire relational network seems unaffected by the intrinsic properties. If one removed the intrinsic phenomenal properties and replaced them with some other intrinsic property, how would this change the physical causal network of causal relations between the events? The Davidsonian move might suggest something like this: The causal relation holds between individual events (not between types of events), and all properties of the event comprise its individual essence. Then causal relations could be affected by changing even seemingly irrelevant properties of a causing event. In the case of panpsychism: If the cause had different intrinsic properties, then the effect would be different. Thus the intrinsic properties are causally efficacious after all.

Sosa criticized this move with a now well-known example: A loud shot kills someone (Sosa 1984, 277). The loudness is epiphenomenal with regard to the killing. Had the gun been equipped with a silencer, the shot would have been equally fatal. In the same sense, mental properties are causally irrelevant within the Davidsonian framework. Davidson replied as follows: "Had the gun been equipped with a silencer, a quiet shot, if aimed as the fatal shot was, and otherwise relevantly similar, would no doubt have resulted in a death. But it would not have been the same shot as the fatal shot, nor could the death it caused have been the same death" (Davidson 1993, 17). A panpsychist might indeed claim that in a world with phenomenal or protophenomenal intrinsic natures, the relata of the causal relations are different from those in a world without such intrinsic natures. Something else is doing the causing and the effect is a different one; therefore the intrinsic properties are causally efficacious after all. But are the causal relations and the causal laws affected by a difference in intrinsic nature? It seems not. In an alternative zombie-world in which nonexperiential intrinsic properties carry the network of causal relations, this network is, on the assumption of the thought experiment, an isomorphic and indistinguishable copy of the causal network in our world. In this world something else is doing the causing, but the causal laws are the same. How are the assumed intrinsic properties of our world causally efficacious? They play a metaphysical role in carrying the causal relations, but they do not really determine those relations. Unless the metaphysical nature of the relata determines the causal relations themselves, the causal efficacy of intrinsic natures remains dubious. They are epiphenomenal with regard to the causal network and the causal laws.

Supervenient Causation?

The strong supervenience relation posited by the constitutive panpsychist could possibly do some work to secure the causal relevance of the mental. In the metaphysical framework of constitutive panpsychism, all causal relations at the macrolevel supervene logically on physical relations at the microlevel. Physical causation relates physical entities. If there are intrinsic properties of the physical, then it is instantiations of these properties that are related by physical causation. The intrinsic properties are mental in some sense (that is, they are microphenomenal). Because the phenomenal properties of the human agent supervene strongly on those microphenomenal properties, the former properties inherit the causal relevance of the latter (Chalmers 1996,154). But what exactly is the causal relevance of these intrinsic properties? They are certainly not needed in scientific causal explanations.

Causal laws operate with physical properties only. Science abstracts away from the quiddities. Nonreductive physicalism is sometimes seen as arguing for dualism with regard to a duality of useful conceptual frameworks but for monism in metaphysics: Even though reality ultimately consists exclusively of physical entities, dualistic talk about the causal role of propositional attitudes and even of phenomenal properties is pragmatically indispensable. Constitutive panpsychism argues in a parallel, but inverted way: Even though reality ultimately has an intrinsic/extrinsic and thus mental/physical duality, it is pragmatically indispensible for science to limit itself to a monism which focuses on the relational physical structure only. But then there is no causal relevance for those intrinsic properties in our practice of explanation. If we abstract from the intrinsic properties, we have at the basic level a net of causal relations which are all physical in the narrow sense. The quiddities are relevant 'only' in a metaphysical sense.

But there are metaphysical worries as well. Does the concept of 'supervenient causation' really make sense? All higher-level causal relations in the world supervene logically on the basic physical relations. What could it mean that the former 'inherit' causal efficacy? The higher-level causal relations are completely microdetermined. All the causal work occurs at the base level, and the higher levels continue to enjoy a metaphysical free lunch. The higher levels *qua* higher levels have no causal efficacy whatsoever. The concept of 'supervenient causation' is misleading, since the supervenient level is asymmetrically dependent on the subvenient base. The causal work is completed at the subvenient level, leaving no causal work to be done by the supervenient level.
'Inheritance' is a misleading concept here, suggesting that the higher level inherits some good from the lower level, with which it could then do something. The supervenient level does nothing at all. It is completely and asymmetrically dependent. The case of logical supervenience might be special. Logical supervenience (across all possible worlds) is almost as strong as identity. A type-identity between higher-level causal structures and the basic physical causal relations would suffice to carry the causal power all the way up. Identity is symmetrical, after all. If the macrolevel is identical with the microlevel, then the levels cannot compete for causal efficacy because no two things that are identical can compete for anything. Analogously, the strong modal force of the relation of strong supervenience might carry causal efficacy all the way up. But mental causation, as many understand it, should make sense of us as agents. If all the causal power is already located in the microconstituents, is there any room left for a robust sense of agency?

But even if we granted a monopoly on causal power to the microconstituents, the problem for the constitutive panpsychist does not disappear entirely.

If the analysis of mental causation requires an account of the causal role of mental properties, then the Davidsonian move is ultimately futile. The basic structure of the problem is this (where 'physical' is used in the narrow sense):

(1) Phenomenal facts do not logically supervene on physical facts.
(2) Facts about causal relations do logically supervene on the physical facts.
(3) Thus phenomenal facts cannot be constitutive for causal relations.

The only way out, it seems, would be to allow for mental-to-mental causal relations. Imagine that, in a cellular automaton, there exist causal laws of the following kind:

> If in a given structural pattern, say, five adjacent cells have the intrinsic property of the type 'hot,' then in the next step the intrinsic property 'pain' will be activated in some cells. There are causal relations from intrinsic states of the cells to future intrinsic states. The relational properties of the cells remain unaffected by these changes.

The problem is that these changes are disconnected from the relational structure of the cellular grid. Since they have no effects on the relational structure, they form a separate and independent causal realm. In effect, this position will have to advocate some kind of psycho-physical parallelism that is somewhat reminiscent of Leibniz's position. But how could such parallelism preserve a robust account of agency in a physical world? What we wish to preserve and account for in any theory of mental causation is not primarily an abstract metaphysics of the causal efficacy of mental properties. It is rather our strong common sense intuition that agency requires the causal efficacy of a higher-level entity as such. Agency requires that my actions as a person are the result of me as a higher-level unity. If the microparticles that constitute me do all the causal work, then the sense of agency is lost. Constitution cannot account for the causal efficacy of the entities constituted at a higher level.

The question is whether the nonconstitutive panpsychist has a better answer to this question. A key problem for the constitutive panpsychist is the ontological status of higher-level mental unities. The panpsychist conceives of the human mind, for example, as constituted by a large number of smaller minds. The so-called combination problem surfaces as a 'composition problem.' Why does an appropriate arrangement of those smaller minds necessitate these higher-level unities? This seems to be a case of strong emergence, since it is conceivable that there is another possible world with the same microlevel mental entities in which no macrolevel mental entities supervene. Constitutive panpsychism must exclude this possibility. The higher-level mind must supervene logically on the appropriately configured microlevel minds.

We have no clue what the compositional principles are that allow this seemingly magical binding of experiences into larger units. We do not understand how subjects sum. This makes us inclined to believe that this special composition question can only be answered by resorting to some form of strong emergence. But the fact that we do not (yet) understand this does not mean that it might not in fact be how nature works. This is how it works everywhere else, the constitutive panpsychist claims. Isn't it then reasonable to assume that the mental as a natural phenomenon is no exception (that is, to assume that nature is homogenous)?

But if this much is granted, another problem looms. According to the constitutive panpsychist, the most basic entities — the furniture of the universe — are experiencers endowed with a point of view or some form of subjectivity. Everything else is made up of these entities. But if this is so, then animals or human experiencers are also basic entities, because they are clearly experiencers and subjects in that sense. Yet the idea of the constitutive panpsychist was that the irreducibly basic entities are only found at the microlevel of nature and that everything else is a composite of them. But if higher-level subjects emerge over the course of evolutionary development, then they belong by definition to the irreducibly basic entities of the universe. Similarly, a proponent of substance ontology can claim that only the basic particles at the microlevel are the true substances making up the furniture of the universe: Carving nature at its joints means describing it at this level. Everything else is just a configuration of these building blocks. In this case a human being is not a substance but a mere configuration of substances. If, however, human beings are counted as substances, then they are ontologically quite independent and not mere configurations of smaller substances. This means that they must be counted as primitive nonreducible particulars in the furniture of the universe. This thought can also be applied *mutatis mutandis* to constitutive panpsychism. The unity of a higher-level subject of experience entitles it to be counted among the basic entities of the universe simply by being a subject of experience. It can then avoid being construed as a mere composition of smaller subjects of experience.

Nonconstitutive Panpsychism, Emergence, and Mental Causation

The nonconstitutive panpsychist bites this bullet. The special composition question is solved by assuming the strong emergence of higher-level mental unities, if the structural conditions at the lower level are met. But this move seems to siphon the theoretical elegance and beauty out of panpsychism. If something like the human mind (or some other animal's mind) does not logically supervene on microphenomenal properties, then we might as well be emergent dualists.

Emergent dualism requires the strong emergence of the phenomenal from the nonphenomenal. Nonconstitutive panpsychism requires the strong emergence of macrophenomenal subjects of experience from microphenomenal subjects of experience. Wasn't the very motive of panpsychism to avoid such a strong emergence? Constitutive panpsychism requires only weak emergence. That is the crucial question that must ultimately be answered. But before we get there, we must first provide a sketch of the metaphysical picture of nonconstitutive panpsychism.

The key question, of course, is which notion of strong emergence the nonconstitutive panpsychist will employ. Even strong emergence comes in different flavors. The main difference in construing emergence seems to be between theories based on nomological supervenience (see Kim 1999) and theories which regard emergence as a nonsupervening causal relation (see O'Connor 2000). The latter is usually taken to provide a more solid grounding for higher-level individuals with novel causal powers (i.e. downward causation). Supervenience construed as an asymmetrical dependency of the higher level on the lower ones seems to preclude any real causal efficacy on the part of the higher levels. The second notion, emergence as nonsupervening causal relation, turns out to entail an even stronger emergence claim. It does this by breaking the bond of synchronic supervenience. If the nonconstitutive panpsychist takes this latter route, then the position will eventually collapse into emergent dualism — or so it seems. But a crucial and decisive difference will remain between the emergent dualist and the constitutive panpsychist. The emergent dualist can solve neither the ultimate-carrier problem nor the problem of the emergence of the phenomenal mind out of entirely nonphenomenal constituents. Nonconstitutive panpsychism, by contrast, has an answer to both of these questions and is thus clearly conceptually different from emergent dualism. Thus, even the version of panpsychism that is closest to classical dualism does not collapse into this well-established position but remains a genuinely distinct account in the philosophy of mind. In the remaining sections this account will briefly be sketched out.

Causal Supervenience

O'Connor and Wong argue that, if emergent entities are metaphysically primitive rather than mere constitutive resultants of lower-level features, then the correct relation between the lower and higher levels is causal, not supervenient (O'Connor and Wong 2005). If these causal connections are indeterministic then we can escape the grip of microdeterminism. For in this case, fixing the microphysical state of the universe will not suffice to fix the distribution of emergent properties—even if we take emergence laws into account. If these laws are indeterministic, then a given microphysical state may have more than one emergent outcome.

Adding temporal dynamics to this picture yields a causal process-oriented metaphysics instead of a formal and static view. The importance of this move cannot be overestimated. A historical example for it would again be Whitehead's "Process and Reality" (Whitehead PR). His classic statement "The many become one and are increased by one" (Whitehead PR, 32) captures the idea succinctly. If the underlying levels of nature reach a certain threshold of complexity of configuration, then an emergent individual is likely to appear. The causal properties of the emergent entity will go beyond the summation of the causal powers at the underlying microlevel. There will be a genuinely new entity with new causal powers. Can this account of emergence handle Kim's worries about causal exclusion (see Kim 1999)? It seems so. On Kim's view, the diachronic causal activities of an emergent entity are metaphysically superfluous. They add no causal efficacy over and above that of the causal mechanisms at the base level. But this contention is based on the asymmetric dependence built into the concept of supervenience. The O'Connor/Wong model is not susceptible to this kind of counterargument. The entities at the basic or subvenient level do not determine the emergent effects independently of the causal activities of those emergent entities (see O'Connor and Wong 2005, 670).

The idea can again be expressed by the analogy of the cellular automaton or 'life world.' Imagine a very complex and long-lasting three-dimensional life world, which mirrors the complexity of our universe. For the longest time, the weak emergence of higher-level structures occurs exactly in accord with the rules governing the individual cells. But occasionally a macroobject appears whose behavior diverges slightly from what the rules predict. The events at the microlevel are affected by this macroobject, changing whenever and wherever such a higher-level structure is present. We would then be able to assign new rules to the situations in which these higher-level structures emerge. It follows from this that the low-level rules do not necessitate the entire future dynamic of the system, not even at the lowest of the hierarchical levels of our three-dimensional cellular automaton. No knowledge of the behavior of those macroobjects or of their effects on the entire system can be derived from knowledge of the smallest microobjects and the rules they follow. If the basic rules are indeterministic such top-down influence might even happen without 'breaking' the most fundamental rules, only the probability distribution will be slightly affected. Such a system is conceptually coherent. If our world were like this, then it would contain strongly emergent entities with downward causal powers. An alternative nonsupervenience view of emergence that would allow for genuine causal efficacy of emergent entities is the fusion account of emergence. Fused entities cease to exist as separate entities. The emergents created by fusion are endowed with new causal powers because the fused entities lose some of their original causal powers (see Humphreys 1997).

Is This Compatible with Science?

Even if these accounts of nonsupervenient emergence are certainly a conceptual possibility, many will claim that we lack empirical evidence for emergent causation. But there are at least candidates for such causal emergence. In mathematical simulations of neural systems it can be shown that the macrolevel can causally supersede the microlevel. This causal independence of the macroobjects, which are more than the sum of the underlying microobjects, is indeed a key feature of Tononi's "integrated information" theory of consciousness (see Hoel, Albantakis and Tononi 2013). Another example may arise in quantum mechanics with the possibility of emerging holistic properties (see Schaffer 2010). Quantum entanglement may be a case in question. Prosser's elegant recent account of this idea can serve as an illustration (see Prosser 2011). Physics does not know of any viable procedure for reducing the entangled state to a summation of classical states and hence reducing quantum mechanics to classical physics. But if that is granted, then the properties of entangled atoms might well be causally efficacious for the future dynamics of the world. Prosser argues: "entanglement shifts a probability distribution concerning the behavior of atoms — the overall configuration of the atoms is likely to be different when there is entanglement — and its effects therefore constitute downward causation. Hence there is no conflict between downward causation by the emergent property of entanglement and the base-level laws" (Prosser 2011, 37). Humphreys's idea of fusion emergence explicitly regards quantum entanglement as involving a fusion of entities into unities of a new kind (see Humphreys 1997). Whether strong emergence in this sense exists is an empirical question. Yet even if the above examples are empirically inadequate, other ways of introducing strong emergence might be found. The mysterious collapse of the wave function itself might be another promising starting point.

Why Nonconstitutive Panpsychism Differs from Emergent Dualism

But why would the panpsychist want to be a strong emergentist? Isn't panpsychism's greatest advantage that it can avoid strong emergence? If higher-level individuals with new causal powers can strongly emerge, then why can't phenomenal minds strongly emerge from an entirely nonphenomenal mindless world? If strongly emergent panpsychism is possible, then so is strongly emergent dualism. And dualism seems intuitively more appealing, since it does not commit us to externally unobservable quiddities or microphenomenal intrinsic properties at the lowest level of nature.

Panpsychism rests ultimately on two theoretical claims: the argument from intrinsic natures and the genetic argument. The argument from intrinsic natures claims that the relational structure described by physics is incomplete. This structure stands in need of categorical intrinsic natures which can carry the relational network of physical properties. As noted above, Leibniz made this point against Descartes, and in the twentieth century Russell and Whitehead advanced the same argument. The dualist has no adequate reply to this objection. Dualism lacks a metaphysically plausible theory of matter. Idealism, which deflates matter to a 'well-founded phenomenon' (see Leibniz), escapes this problem. But the Cartesian substance dualist faces not just the well-known problems of causal interaction between spatial and nonspatial entities, but also that of developing a plausible metaphysical theory of matter. Mere extension or mere relation will not do the trick. It leaves open the questions of what is being extended and what is being related. Panpsychism overcomes this dearth of analysis. This is the first reason why nonconstitutive panpsychism does not collapse into emergent dualism.

The second reason is related to the genetic argument for panpsychism. Even O'Connor and Wong, who are no panpsychists, admit that their concept of emergence requires a "tendency had by each of the basic entities" (O'Connor and Wong 2005, 665) which explains the strong emergence of phenomenal minds. But what exactly does "tendency" mean here? Can something entirely nonphenomenal and nonmental have the tendency to bring about phenomenal minds? Earlier we introduced the distinction between inter-attribute and intra-attribute emergence. Interattribute emergence might be labeled 'superstrong emergence.' In addition to weak and strong emergence we thus introduce superstrong emergence. Weak emergence (based on strong-supervenience relations) and strong emergence (based on weak-supervenience or causal relations) occurs within a unified categorical framework. For example: Higher-level spatio-temporal concrete entities can only emerge either weakly or strongly from lower-level spatio-temporal concrete entities. Superstrong emergence breaches or transcends categorical frameworks. A clear case of superstrong emergence would arise if something emerged from absolutely nothing. Another clear case of superstrong emergence would arise if a concrete spatio-temporal entity emerged in a world in which only abstract entities exist. The panpsychist claims that another case of superstrong emergence is given by the emergence of the phenomenal mind from a world which is merely spatial extension or a framework of causal-functional relations.

The nonconstitutive panpsychist needs strong but not superstrong emergence. There are new unified entities endowed with phenomenal minds that are more than just the constitutive summation of smaller such entities. But even the smallest entities have some form of phenomenal properties. Phenomenality was there in the beginning. Likewise, the strongly emerging higher-level entities feature new causal powers that are more than just the constitutive summation of the causal powers of smaller such entities. But even the smallest entities have some form of causal properties; causality was there in the beginning. What is new is simply the strong emergence of new natural individuals (concrete entities) that cannot be reduced to their constituents. What strong emergence makes possible is the 'special composition' of new entities from the same basic makeup as their constituents. Strong emergence cannot, however, create something absolutely new in its metaphysical nature. This recalls again to mind Whitehead's classic dictum: "The many become one and are increased by one" (Whitehead PR, 32). Emergent dualism postulates the emergence of a whole new metaphysical category of entities. Mental entities emerge from nonmental entities. This is a case of superstrong emergence as defined above. Emergent dualists thus not only have no answer to the problem of causal pairing and that of the intrinsic natures, they also require superstrong emergence for the mental to emerge from the physical. Thus, panpsychism, even in the strong-emergentist form of nonconstitutive panpsychism, clearly differs from dualism.

Can the nonconstitutive panpsychist get a grip on the problem of mental causation? The O'Connor/Wong model of emergence contains genuine down-ward causation. Higher-level entities endowed with minds could in principle have causal powers that are not microdetermined by the causal relations at lower levels. The nonconstitutive panpsychist therefore need not be burdened with the problem that all macrocausal relations strongly supervene on microcausal relations. There is an opening for macrolevel agents to make a causal difference in the world. This is a genuine difference from constitutive panpsychism.

Some would argue, however, that genuine mental causation requires that the mental content of the psychological states of these entities or agents be causally efficacious. The nonconstitutive panpsychist can block this argument by the very 'Davidsonian move' made above by the constitutive panpsychist. According to the Davidsonian move, it is the entire entity that is doing the causing, and if its intrinsic (mental) properties were to change, something else would do the causing and something else would be caused. But again, this option renders too many properties causally efficacious, as Sosa rightly points out (Sosa 1993). The problem of mental causation seems all but intractable. No position in the philosophy of mind has a fully convincing answer to it. But non-constitutive panpsychism is certainly not in a theoretically precarious position with regard to mental causation.

It has the theoretical advantage over emergent dualism that it does not require a causal pairing of spatial and nonspatial entities. Nonconstitutive panpsychism is again clearly distinct from dualism.

Taking Stock

It has been shown that panpsychism is a conceptually stable and independent position in the philosophy of mind. It differs from physicalism, dualism, idealism, and even neutral monism. It comes in two flavors. One — constitutive panpsychism — incorporates many ideas from type-A physicalism without collapsing into it. The other — nonconstitutive panpsychism — incorporates ideas from emergent dualism without collapsing into it. I admit to having (at least sometimes) some sympathies with Kantian worries about the metaphysical inscrutability of the relation between the mental and the physical. So I am not claiming that a version of panpsychism is true. But I am claiming that it might be. Both versions of panpsychism discussed here are certainly respectable and coherent positions; panpsychism needs to be taken seriously. Thus: "*If it is not true, it is well conceived—Se non è vero, è molto ben trovato.*"

Bibliography

Brüntrup, Godehard. "Natural Individuals and Intrinsic Properties." *Unity and Time in Metaphysics*. Ed. Ludger Honnefelder, Edmund Runggaldier, and Benedikt Schick. Berlin/New York: de Gruyter, 2009. 237–52.

Brüntrup, Godehard. "Panpsychism and Structural Realism." *New Perspectives on Panpsychism*. Ed. Michael Blamauer. Frankfurt: Ontos, 2011. 15–35.

Chalmers, David. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press, 1996.

Davidson, Donald. "Thinking Causes." *Mental Causation*. Ed. John Heil and Alfred Mele. Oxford: Oxford University Press, 1993. 3–18.

Haugeland, John. "Pattern and Being." *Dennett and his Critics*. Ed. Bo Dahlbom. Oxford: Blackwell, 1993. 53–69.

Hoel, Erik, Larissa Albantakis, and Giulio Tononi. "Quantifying Causal Emergence Shows that Macro Can Beat Micro." *Proceedings of the National Academy of Sciences* 110.49 (2013): 19790–795.

Humphreys, Paul. "How Properties Emerge," *Philosophy of Science* 64.1 (1997): 1-17.

James, William. *The Principles of Psychology*. Cambridge: Harvard University Press, 1983.

Kant, Immanuel. *Critique of Pure Reason*. Transl. Norman Kemp-Smith. London: Macmillan, 1929.

Kim, Jaegwon. "Making Sense of Emergence." *Philosophical Studies* 95.1–2 (1999): 3-36.

Leibniz, Gottfried. *Die philosophischen Schriften*. Vol. I–VII. Hildesheim: Georg Olms, 1965.

Newman, Max. "Mr. Russell's Causal Theory of Perception." *Mind* 37.146 (1928): 137–48.

O'Connor, Timothy. *Persons and Causes: The Metaphysics of Free Will*. Oxford: Oxford Universiy Press, 2000.

O'Connor, Timothy and Hong Yu Wong. "The Metaphysics of Emergence." *Nous* 39.4 (2005): 658–78.

Prosser, Simon. "Emergent Causation." *Philosophical Studies* 159.1 (2011): 21–39.

Putnam, Hilary. "Models and Reality." *Journal of Symbolic Logic* 45.3 (1980): 464–82.

Rosenberg, Gregg. *A Place for Consciousness. Probing the Deep Structure of the Natural World.* Oxford: Oxford University Press, 2004.

Russell, Bertrand. *Our Knowledge of the External World as a Field for Scientific Method in Philosophy*. Chicago: Open Court, 1914.

Russell, Bertrand. *The Analysis of Matter.* London: Routledge, 1927.

Schaffer, Jonathan. " Monism. The Priority of the Whole." *Philosophical Review* 119.1 (2010): 31–76.

Sosa, Ernest. "Mind-Body Interaction and Supervenient Causation." *Midwest Studies in Philosophy* 9.1 (1984): 271–81.

Sosa, Ernest. "Davidson's Thinking Causes." *Mental Causation*. Ed. John Heil and Alfred Mele. Oxford: Oxford University Press, 1993.

Whitehead, Alfred. *Process and Reality: An Essay in Cosmology*. Corr. ed. Ed. David Griffin and Donald Sherburne. New York: Free Press, 1978.