Ida Skovhus Hansen:

Strong Emergence and its Relation to Philosophy of Mind

RESUMÉ

Denne artikel forsøger at besvare spørgsmålet om hvorvidt stærk emergens er et holdbart begreb, hvad det medfører, om der findes eksempler på stærk emergens og hvordan dette vedrører bevidsthedsfilosofi. Ifølge stærk emergens er visse sandheder om et givet høj-niveaus fænomen ikke reducerbare, end ikke principielt, til det fænomen på lavere niveau som det afhænger af. I artiklen argumenteres der for at selvom stærk emergens er et holdbart begreb, så findes der ikke nogen overbevisende og uproblematiske eksempler i naturen. Ydermere medfører stærk emergens kontraintuitive konsekvenser for kausale strukturer og bevidsthedsfilosofi.

ABSTRACT

This paper attempts to answer the question of whether strong emergence is a tenable concept and examine how it relates to philosophy of mind. Strong emergence is the idea that truths about a given high-level phenomenon are not reducible, even in principle, to the low-level phenomenon on which it depends. This article advances the position that strong emergence, though a tenable concept, cannot present any convincing and unproblematic instantiations in nature. Furthermore, strong emergence might imply counterintuitive consequences for causational structures and philosophy of mind.

EMNEORD

metafysik, bevidsthedsfilosofi, emergens, reduktion

KEYWORDS

Metaphysics, philosophy of mind, emergence, reduction

Introduction

In this paper, I will discuss strong emergence in relation to philosophy of mind, examine whether or not we can find any examples of strong emergence, and explore what the consequences of the position are. I will argue for an anti-strong emergence stance and use this to exclude certain possibilities within philosophy of mind. I will start by defining emergentism and exploring the coherence and consequences of the definitions of emergence. I will place a special focus on strong emergence, as this will turn out to be especially problematic. I will examine cases where emergence is invoked frequently, including those involving emergence in philosophy of mind. Finally, I will use my anti-strong emergence stance to exclude certain possible answers to the mind-body problem.

Defining Emergence

Many scientists and philosophers do not distinguish between strong and weak emergence, although these are two separate concepts. This is because the concept of emergence is often used in many different contexts and ascribed different meanings. When the concept is employed to describe a situation in science and philosophy, confusion often arises as to the kind of emergence meant. Here, I will start by differentiating between the two different concepts. Strong emergence is defined by David Chalmers (2006, 244) as follows: "a high-level phenomenon is strongly emergent with respect to a low-level domain when the high-level phenomenon arises from the low-level domain, but truths concerning that phenomenon are not deducible even in principle from truths in the low-level domain." According to the vitalists, life can be seen as an example of strong emergence, where something extra is needed in order to explain how life arises from lifeless matter. Weak emergence, on the other hand, is defined as follows: "a high-level phenomenon is weakly emergent with respect to a low-level domain when the high-level phenomenon arises from the low-level domain, but truths concerning that phenomenon are unexpected given the principles governing the low-level domain" (ibid.). An example of such a phenomenon is wave propagation in fluids; this behavior may be unexpected, but it is nonetheless describable in physical terms. In general, weak emergence applies to emergent phenomena as follows: "any emergent

phenomenon, say Y, is wholly dependent on that which it emerges from, say X" (Strawson 2006a, 14).

These definitions assume a layered view of reality, in which we have high- and low-level phenomena, objects, etc. This layered view of reality raises many further questions: where can levels be identified, are there definite boundaries between them, what kind of relations connect them, etc.? In this paper, I will assume a layered view of reality akin to the divisions between the sciences and thus expect a rising level of complexity, as we work from physics upwards (Ellis 2006, 80). I will return to the relationships between the sciences later in the paper. However, different models of layeredness can be assumed, and the argument of the paper will still apply.

From the definitions of strong and weak emergence respectively, we can deduce that strong emergence often implies weak emergence, depending on the interpretation of the word "unexpected" (Chalmers 2006, 245), but cases of weak emergence need not imply strong emergence.

We can also conclude that weak emergence is relative to the observer (Chalmers 2006, 251), as it depends on the individual's interpretation of "unexpected". Weakly emergent properties are properties arising at a high level, which are not easy to deduce from low-level properties but are nevertheless deducible from them. Chalmers (2006, 252), however, worries that we might include many phenomena not classically considered weakly emergent if we think of "unexpectedness" in too wide a sense. That is, we can ask ourselves whether it really is unexpected that complex phenomena arise from low-level properties structured in complex ways, or whether we should restrict "unexpectedness" to cases in which high-level phenomena emerge from low-level phenomena structured in simple ways.

Given the many different ways of construing weak emergence depending on the way we look at a phenomenon, we may well see it as a cluster concept, in which different subtle definitions all point toward some common features or some paradigm examples (Chalmers 2006, 253). The important point here is the crucial difference between weak and strong emergence; strong emergence is not relative to an observer but instead constitutes an ontologically independent phenomenon.

Weak emergence, in spite of its observer-dependence, may still assist us in understanding the nature of our reality. Weak emergence can help us get a better understanding of complex phenomena; it is therefore still very important when we describe the world in scientific terms. In this way, it might help make a physicalist worldview more plausible:

> by showing how a simple starting point can have unexpected consequences, the existence of weakly emergent phenomena can be seen as showing that an ultimately physicalist picture of the world need not be overly reductionist, but rather can accommodate all sorts of unexpected richness at higher levels, as long as explanations are given at the appropriate level. (Chalmers 2006, 246)

In opposition to this, thinking along same the lines as Galen Strawson (2006a) and Chalmers (2006, 246), we might say that if strong emergence exists, it will have the potential to reject a physicalist worldview. A physicalist worldview is one in which causes have determinate effects. Strong emergence will introduce a sort of indeterminism, where we can never know which higher phenomena arise from lower ones or how they do so;l it will make certain matters brute facts, things that we should just accept (Kim 2011, 305; Nagel 1979, 187).

Strawson (2006a, 12) also questions the coherence of strong emergence altogether. Yet even if the idea of strong emergence is coherent, as Jaegwon Kim (1999, 6-7) concludes that it is, we still need to find out whether cases of strong emergence actually exist. Kim adopts a view of strong emergence as including an inexplicability and a sort of unpredictability. To see this, he first assumes that systems can in principle be given a complete micro-structural description, including constituents, intrinsic properties of the constituents, and the relations between these. Next, he assumes mereological supervenience, defined as such: "Systems with an identical total micro-structural property have all other properties in common. Equivalently, all properties of a physical system supervene on, or are determined by, its total microstructural property" (Kim 1999, 7). Both of these assumptions are related to the causal closure of the physical-- that is, by the assumption that all physical effects have sufficient physical causes and by a commitment to reduction. Among the supervening properties, some are emergent, whereas others are simply resultant. Where supervenience can be defined as a relation in which if A is supervenient on B, this means that whenever B occurs, A must also occur; that is, there can be no

change in A without a change in B (Ney 2014, 56). There might, however, also be cases where the supervenience of the strongly emergent phenomenon on the base fails. These will be cases where the causal relation of the phenomenon upon the base is indeterministic (O'Conner 2020).

Kim (1999, 8) identifies the main feature making some properties strongly emergent as being not theoretically predictable from the micro-structural properties of a system. Strongly emergent properties are thus inductively predictable. That is, if a system exhibits a strongly emergent property at a certain time, *t*, and retains its micro-physical properties, then we can, inductively, predict that the given system will exhibit the same strongly emergent property at a later time, t_2. But according to strong emergentists, we cannot theoretically predict a strongly emergent property, even given full knowledge of a system. This theoretical unpredictability means that in the case of phenomenal properties, strong emergence can only be recognized by experience, as it cannot be predicted on the basis of any information about the underlying system. In these cases, such an emergent property is simply a brute fact for the strong emergentist; there will be no explanation of why the property is correlated with a specific micro-structure. It also has the consequence: if we were to design something with phenomenal properties, we could only do so on the basis of inductive prediction; that is, we could only assume that something with the exact same constitution as something that we know has these properties (for example, us) would also exhibit phenomenal properties. As we have no theoretical basis for explaining why or how phenomenal properties occur, theoretical and novel predictions of phenomenal properties would not be possible.

Likewise, the strong emergentist would not be able to make an ontological reduction. An ontological reduction presupposes that when we reduce from high-level (something complex) to something low-level (something simple), we are left with a simpler ontology, because it posits fewer elements. But in the case of strongly emergent properties, an ontological reduction is not possible, as we cannot reduce the properties to a simpler base. Instead, the strongly emergent property constitutes a novel addition to our ontology. This kind of ontological reduction need not, but can of course, remove the high-level property from our ontology. It can, for example, conserve it as something

simpler or as part of something simpler, for example via an identity relation (Kim 1999, 15).

As a consequence of strongly emergent properties being both theoretically unpredictable and nonreductive, they become scientifically inept. We can never theoretically predict their existence or their causal influence, if they have any, nor can we explain their existence in the form of a reduction of any kind. They become something we just have to accept the existence of, which can never be explained further or used for anything besides inductive predictions. They must, as put forth by the early emergentists, be accepted with a "natural piety"(Alexander, 1920, 47; Morgan, 1923, 5-6; O'Connor and Wong, 2020).

Having clarified the inexplicability and unpredictability of strong emergence, we can see why Strawson (2006a, 18) believes the notion to be incoherent. The lack of intelligibility of strong emergence in a metaphysical sense – that is, the fact that there is "absolutely no reason in the nature of things why the emerging thing is as it is (so that it is unintelligible even to God)" – is exactly why Strawson rejects the notion as incoherent. Strong emergence becomes a miracle every time it occurs, because there is nothing about the base that gives reason for the emerging phenomenon to emerge. Even if the strong emergentist assumes that a supervenience relation exists between what emerges and the base it emerges from, this will only amount to a law-like miracle (ibid.). Furthermore, we will have to show in cases of suspected strong emergence that this is the only or the best option for explaining a given phenomenon, in order to avoid adding superfluous elements to our ontology. Strong emergence can thus seem metaphysically extravagant (Strawson 2016, 83). Applying the principle of Occam's razor, encouraging a lean ontology, Strawson excludes strong emergence if no actual cases can be found and it is not shown to be necessary in order to explain phenomena in metaphysics.

The strong emergentist might object at this point that it seems that we find other brute facts in nature, e.g., the basic laws and properties in physics (ibid.; Chalmers 2002, 262). These, however, are posited as simply fundamental. This need not be a problem. According to Strawson (2006a, 18), it is allowable that there is "no reason for it in the nature of things" when we talk of something at the most fundamental level; when something emerges, however, it is wholly dependent upon a base but still cannot be explained or necessitated from that base. If we had brute emergence in this way, where something with a different

nature from its base, on which it is wholly dependent, arises, nothing could be ruled out. We would leave the determinate world of science, and indeterminism would be hard to exclude, as we would have a phenomenon for which there could be no physical explanation. That is, we would have no way to theoretically predict it; indeterminism would therefore, at the very least, be hard to exclude in the case of such a phenomenon.

Now that we have characterized strong and weak emergence and identified the problems which seem to follow in the wake of the notion of strong emergence, we still need to ask whether any cases of strong emergence actually exist and, if such cases are found, how they construe of the workings of the emergent phenomenon. In the next section, I will be exploring these questions.

Strong Emergence?

We can now ask: do strongly emergent phenomena actually exist? When we ask such a question, what we are really asking is: is there anything besides the physical and the things which can be reduced to this, including things which are weakly emergent?

Many of the classical examples put forth as examples of strong emergence turn out to be weakly emergent instead. This is the case for life from lifeless matter (Strawson 2006a, 20), which the vitalists earlier conceived of as a mystery. It is a case of weak emergence from biological and chemical phenomena, as it does not need to be seen as a wholly new phenomenon in order to be explained. It can, in other words, be reduced, even though this reduction is not a simple one. The same goes for the case of evolution, which is also weakly emergent. Here we find the gene as an underlying base, which is itself complex, but from it emerges something unexpected with an even higher degree of complexity--namely the evolutionary process when combined with mutations, combinations with other genes, and environmental constraints (Chalmers 2006, 251). Another case of weak emergence which has sometimes been argued as a case of strong emergence is the phenomenon of heat or temperature, which can be reduced to the mean kinetic energy of moving particles (Kim 1999, 11) and therefore exemplifies weak emergence as well. More recent examples are often found in computer science, where connectionist networks (Chalmers 2006, 252) and other complex phenomena which arise out of simple rules set up in a

program (Ellis 2006, 81) may be unexpected but can still be reduced to the rules of the program and the algorithm and data fed to it.

The consequences of strong emergence are perhaps most clearly articulated in physics:

> If there are phenomena that are strongly emergent with respect to the domain of physics, then our conception of nature needs to be expanded to accommodate them. That is, if there are phenomena whose existence is not deducible from the facts about the exact distribution of particles and fields throughout space and time (along with the laws of physics), then this suggests that new fundamental laws of nature are needed to explain these phenomena.
> (Chalmers 2006, 245)

When we talk of physics, clarification is needed. A new discovery by physicists leading to a new fundamental law might at first sight seem to qualify as a strongly emergent phenomenon. This, however, is not the case. In this paper, it will suffice to subscribe to the definition of the physical given by Thomas Nagel (1979, 183): we can count as physical all that can be discovered by explanatory inference from what is already within the field of physics.

As soon as we move to higher levels of complexity than the ones we find in physics, we need only consider whether the phenomena we encounter are in principle deducible from a physical description of the world. That is, we need not have a full description in physical or functionalist terms. We might say, "no new fundamental laws or properties are needed: everything will still be a consequence of physics" (Chalmers 2006, 245). Likewise, Kim (1999, 18) also asserts that these higher-level phenomena, although not clearly reducible, all seem to be functionalizable and given a physical base. This makes them reducible; they might of course still be weakly emergent, with the possible exception of qualia. This is not the case with our current physics (Ellis 2006, 102), where we do not have the means to describe most higher-level phenomena in terms of some underlying reduction; it is far from clear, however, that any higher-level phenomenon, besides perhaps phenomenal consciousness (what we might call experience or what-is-it-likeness), can be said to involve novel laws of nature. Due to the weakly emergent nature of the cases offered by many strong emergentists, several philosophers conclude that there might just be one

case where strong emergence possibly exists: the case of consciousness. Kim (1999, 18) states that "if anything is going to be emergent, the phenomenal properties of consciousness, or 'qualia', are the most promising candidates." Chalmers (2006, 247) takes a similar stance, claiming that "I think that there are no other [than consciousness] clear cases, and that there are fairly good reasons to think that there are no other cases".

The bleak outlook for strong emergence does not mean that higher-level phenomena cannot be unexpected given lower-level phenomena (weakly emergent); it does, however, mean that none of these higher-level phenomena involves entirely new conceptions of nature, with the possible exception of the certain quality of consciousness.

Picking up on the assertion of a layered view of reality, we can question whether strong emergence relations exist between the different levels or whether these too can be reduced to the lowest level, retaining weak emergence as a possibility. Following Chalmers's earlier definitions of strong and weak emergence, we might wonder whether high-level laws are deducible from low-level laws alone. Chalmers's (2006, 248) reply to this invokes low-level facts; if high-level laws are not deducible from low-level laws alone, they might still be deducible from a combination of low-level laws and low-level facts. If this is the case, it implies that low-level facts will not follow from low-level laws alone either; they would then be deducible from these low-level laws, but something further would still be needed. This would be a case in which we must add something to our physical laws to be able to fully describe higher-level laws. The higher-level laws thus have an influence on what we need to posit on a lower level. Several philosophers turn to causal explanations of these phenomena, mostly in relation to consciousness (Chalmers, 2006; Kim, 2011; Nagel, 1979). Chalmers describes the situation as involving a sort of downward causation, which he defines as follows:

> Downward causation means that higher-level phenomena are not only irreducible but also exert a causal efficacy of some sort. Such causation requires the formulation of basic principles which state that when certain high-level configurations occur, certain consequences will follow [...]. These consequences will themselves either be cast in low-level terms, or will be cast in high-level terms that put strong constraints on low-level facts. (2006, 248)

Downward causation can be distinguished from same-level causation and upward causation (Ellis, 2006, 82; Kim, 1999, 22). Same-level causation is causal action at the same level of complexity, for example between brain states; upward causation is causal action from a lower level to a higher level of reality, for example brain states causing mental states.

Like emergence, downward causation comes in two flavors: a strong and a weak. In strong downward causation, the causal influence of the high-level phenomenon on a lower-level phenomenon is not deducible even in principle from low-level laws and initial conditions; if we were to find a case of strong downward causation, this would constitute an example of strong emergence. With weak downward causation, the high-level phenomenon is unexpected based on low-level laws and initial conditions, but it is still deducible in principle (Chalmers 2006, 249). Kim (1999, 26) supplies us with an example of weak downward causation: a vase being thrown from a high window. In this case the mass and speed of the vase will influence the air molecules around it on its way down. Here the total mass of the vase (a high-level object), and not just the parts of it, will have an effect on the air molecules (lower-level objects) surrounding it. In this way, we can find many examples of differing scales of weak downward causation, but it is doubtful that we will find strong downward causation anywhere.

It might be the case that strong downward causation occurs in collapse interpretations of quantum mechanics (Chalmers 2006, 249). According to quantum mechanics, a Schrödinger wave function is the description of a given particle/wave. The wave function gives a deterministic description of the particle/wave development but can on occasion experience an indeterministic collapse. These collapses occur when measurements are performed (Chalmers 2002, 262). There is, however, no good definition of measurement in physics, especially in the field of quantum mechanics. The lack of a definition opens several possibilities, one of which being a model of strong downward causation. Here the collapse can be interpreted as a result of the influence of the high-level phenomenon of measurement, interpreted as involving a conscious observer, which forces the wave function to behave in ways it otherwise would not. In this case the high-level phenomenon of measurement by a conscious observer will exert strong downward causation on the low-level phenomenon of the wave function. The collapse interpretation of quantum

mechanics can be seen as an example of strong downward causation and also of strong emergence, because it involves new fundamental laws and these are in principle non-deducible. However, in this case we have to remember that there are several interpretations of quantum mechanics. Furthermore, there is no consensus in physics as to which interpretation is the correct one, even though the differing interpretations have widely different consequences both in physics and philosophy.

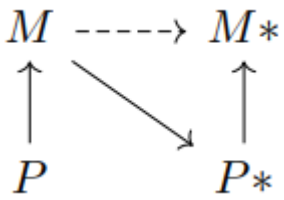
The case of the interactionist interpretation of the collapse of the wave function in quantum mechanics is quite different from the case of consciousness, which some see as another place where strong downward causation can be found. Consciousness itself can be seen as a strongly emergent phenomenon and need not involve any form of strong downward causation. If, however, consciousness as an emergent quality did not involve strong downward causation, this would point us in the direction of some kind of epipheno-menalism. According to this view, the high-level phenomenon would not be able to have any causal influence on low-level phenomena, including brain states, as the mental states of consciousness would supervene upon these. If such an epiphenomenalism is adopted, we only have upward causation from brain states to the emergent mental states. The two can, however, also be combined. If we combine a strongly emergent view of consciousness with a downward causation, then consciousness will have a causal influence on low-level phenomena. Chalmers (2002, 262) combines this with the aforementioned interpretation of quantum mechanics as he interprets measurement in a way in which consciousness plays a crucial role. Consciousness therefore becomes instrumental in the collapse of the wave function.

Most emergentists who subscribe to the view that consciousness is strongly emergent also opt for a version where the strongly emergent phenomenal property exhibits causal downward influence, instead of epiphenomenalism (Alexander, 1920, 8; O'Connor and Wong, 2020). They do, however, also acknowledge that brain states determine mental phenomena; therefore, the situation becomes more complex. The strong emergentist is left with a picture in which brain states control mental phenomena and mental phenomena affect brain states. This must be the case for most of the emergen-tists, as their thesis often relies on the supervenience of emergent phenomena from lower-level phenomena. The lower-level phenomena therefore determine the existence of

the emergent phenomena altogether. Yet a problem seems to arise if we assume that all this causal action happens at the same time, *t*, leading to a causal circularity (Kim 1999, 29). In order to avoid such a causal circularity, we must assume that the causal interactions between the emergent pheno-menon and the low-level phenomena from which it arises happen at different times, although we can allow this time difference to be minimal.

If we apply this to Chalmers's (2002, 263) collapse interpretations of quantum mechanics, it actually makes a measurable difference. Given instruments sensitive enough, we should be able to measure a difference in time, even if it is miniscule; this renders it testable whether or not a case can be made for there being strong downward causation in quantum mechanics.

However, Kim (1999, 32) finds that despite the option of allowing a time difference between the effects of the phenomena on each other, a further problem arises that makes the position even harder to uphold. If we have a physical condition *P*, which causes a strongly emergent mental phenomenon, *M*, then in order for *M* to cause a same-level phenomenon, it is presupposed that *M* exerts downward causation and influences a physical base *P**, which then causes *M**, the same level phenomenon, which can but need not be an emergent property. We can represent this in a diagram:



P is the physical base, M the mental state, M the resulting mental state, and P* the physical base needed to constitute M*. The arrows represent the causal action—their direction tells us whether it is upward, same-level, or downward causation.*

But if we understand causation as nomological sufficiency, *P* becomes nomologically sufficient for *P**, thus rendering *M* superfluous. *P* becomes the condition without which *P** would not have occurred. Further, *M* cannot be viewed as a link in a causal chain between *P* and *P**, as "the emergence relation from *P* to *M* cannot properly be viewed as causal" (ibid.). This results in *M* being causally superfluous, or in the case of *M* being mental phenomena, a form

of epiphenomenalism. This would not be the case if instead of a strongly emergent relation we had a causal relation of some sort, making *M* either weakly emergent and thereby unexpectedly reducible to its low-level base or simply reducible.

The rejection of strong downward causation, and thereby also the possibility of holding a strongly emergent view that does not entail some form of epiphenomenalism, is therefore not only facilitated by the assumption of the causal closure principle of physics but also by the fact that the emergence relation cannot be properly viewed as a causal relation. In order to prove that strong downward causation could work and give the strongly emergent phenomenon the opportunity to causally influence lower-levels of reality, the emergentist would need to supply a positive argument in favor of the position or deny the causal closure of the physical.

We might encounter the objection that causality itself, and especially mental causality, is highly problematic. Additionally, in cases in which strong emergence is not assumed, we find muttered and vague notions of mental causality, and the causal closure of physics leads to the problem of overdetermination for most kinds of theories of consciousness. However, it seems that strong emergence poses a special problem, as the relation between the emergent phenomenon and its base is not even causal.

Having examined the problematic definitions of strong emergence, potential cases of strong emergence, and their possibility of causal interaction with lower-level phenomena, I argue that the case for strong emergence is weak and lacks a positive argument in favor of the position or real examples to use as models. Following this, clarification is needed as to the consequences of rejecting strong emergence. In the following section, I will focus on where this rejection leads with regard to positions in philosophy of mind, as this is an area in which strong emergence is often invoked.

Where Does This Leave Us?

Which views are left in philosophy of mind when strong emergence is denied? Many of the authors advocating a rejection of strong emergence are led to forms of panpsychism, although this is not the only option. Panpsychism can be described as a view which asserts that everything has some sort of experiential

being, that consciousness is fundamental (Goff and Allen Hermanson 2020). This view has been adopted by both Nagel and Strawson (1979; 2006a), although the forms of panpsychism advocated by these authors differ notably. In this section, I will explore the options left following a denial of strong emergence.

Strawson (2006a, 11-12) characterizes all views accepting the following two statements as strongly emergent:[1] "[NE] physical stuff is, in itself, in its fundamental nature, something wholly and utterly non-experiential," and "[RP] experience is a real concrete phenomenon and every real concrete phenomenon is physical." In order to combine these two theses, some form of strong emergence must be invoked (ibid.), which according to Strawson makes the combination incoherent.

The first statement is often accepted by physicalists,[2] who assume that only physical phenomena exists, along with substance and property dualists, who can be taken to assert either the existence of separate substances of the mental and the physical resulting in substance dualism or the existence of mental properties that cannot be reduced to physical properties (Chalmers 2002, 261). But the statement rules out panpsychism and views akin to this, such as panprotopsychism and what Chalmers calls type-F monism or Russellian monism. Panprotopsychism can be characterized as a version of panpsychism in which the smallest constituents of the universe can be said to have protophenomenal properties or the potential to constitute conscious phenomena (Goff and Allen-Hermanson 2020). Russellian monism is the view that "consciousness is constituted by the intrinsic properties of fundamental physical entities" (Chalmers 2002,265; 2010,133).

The second statement rules out both kinds of dualism, as they both will not agree that every real concrete phenomenon is physical. It also rules out eliminativism, which can be described as a view in which the existence of consciousness and phenomenal truths is denied (Chalmers 2002, 251). The eliminativist position is ruled out by the first part of the second statement, as

[1] Strawson calls strong emergence *radical emergence*.

[2] Here I will use physicalism instead of materialism, which might lead to ideas of the existence of only material objects

the eliminativists do not agree that experience is a real concrete phenomenon, whereas dualism is ruled out by the second part of the statement.

Strawson (2016) opts for rejecting the first statement, leading him to a form of panpsychism. He states that we have no evidence of the physical being nonexperiential and refers to the metaphysical simplicity and problem-solving power of a panpsychist position, where there is no need for bridge-laws between conscious phenomena and the physical. Instead, these would constitute a form of strong emergence, something to be added to our physical laws.

Emergentist views have often been seen as forms of nonreductive physicalism (Kim 1999, 4), where it is postulated that nothing besides the physical exists, but novel nonreducible and nondeducible phenomena arise with a certain level of complexity. Emergentists often see this position as a middle ground between a strictly reductive physicalism and dualism. As argued in this paper, however, this position does not seem tenable, due to its strong emergentist commitments.

However, we might question whether panpsychism actually avoids strong emergence. In *Experiences Don't Sum*, Philip Goff (2006, 53) concludes that the combination problem, which arises for panpsychists, might force the panpsychist into accepting some form of brute emergence. The combination problem can be described as the problem of how a macro-consciousness like ours comes to be constituted from billions of "experience-involving ultimates". It is hard to see how this macro-consciousness should not be a new fact, separate from facts at the low-level, and how it arises from billions of independently experiential ultimates.

We might respond that it is still an option that the gap in knowledge of the constitution of a macro-consciousness is only weakly emergent; that is, it might be in a sense an epistemological gap, a lack of knowledge, and a sense of unexpectedness. Goff (2006, 56) responds that it seems the reductive physicalist is in the same position here as the panpsychist; she must assume that conscious experience emerges at some point, whereas the panpsychist must assume the same about macro-experience. Further, it seems that the reductive physicalist can accept the epistemological gap as a gap in our knowledge, without assuming that this gap has the metaphysical consequence of the physical being fundamentally experiential (ibid.). To this we might object, following Strawson, that it is difficult to see how the position of reductive physicalism escapes

strong emergence, as proponents of the positions ascribe to both NE and RP. One exception might be identity theory, which dissolves the problem of consciousness by assuming that conscious states are identical to certain physical states; for example, pain is identical to c-fibers firing. Thus, identity theorists can be said to deny NE but accept RP. Identity theory, however, encounters objections, most notably arguments based on multiple realizability. These arguments claim that the same mental state (e.g., pain) can be realized in several ways and may not be confined to a single physical realization (e.g., c-fibers firing).

Goff also objects that Strawson makes the hidden assumption that our conscious experience is completely transparent for us. This entails that we would, if we were constituted by billions of experiential ultimates, experience what they experience and not some novel composition of the ultimates' experiences, which would be the case if some form of emergence is not accepted.

In his reply, Strawson (2006b, 250-252) objects that a macro-consciousness constituted from experiential ultimates can be given the analogy of physics. Many macroscopical phenomena in our lives do not seem to be constituted by microscopic particles, but they are nevertheless constituted in this way and all these phenomena are only weakly emergent . Further, Strawson clarifies that he only subscribes to a partial transparency of conscious experience and that the hidden parts may be the nature of the constitution and compositions itself.

It thus seems that the panpsychist position is still open after rejecting strong emergence; however, this works only with the acceptance that the combination problem must be answered in a way which avoids strong emergence if possible, which may be doubted following Goff's objections.

Where does this leave us? If we can find no examples of strong emergence, if the case for it leads to several problems including it being in principle theoretically inexplicable, and if strong downward causation is required in order for a strongly emergent phenomenon to be causally efficacious, it seems plausible to reject the notion of strong emergence altogether. Such a rejection of strong emergence will exclude from philosophy of mind all the views which accept the closure of physics but believe conscious phenomena to not be deducible from these laws or believe that physical stuff is experiential. All views including a commitment to non-reductive physicalism will be ruled out by such

a rejection. We seem to be left with the options of eliminativism, dualism (although some further causal explanations would be needed in this case, and the property dualist might have to become an epiphenomenalist), identity theory, and perhaps some forms of panpsychism, including Russellian monism.

In the current landscape of philosophy of mind, the best option for the strong emergentist to retain her position might be epiphenomenalism in its different flavors. Such a position might seem undesirable, as it renders mental states incapable of interacting with lower-level phenomena. This would be opposed to intuition; for example, it would go against the intuitive idea that my wanting to go to get coffee has something to do with me going to get coffee, as it would render my intention to get coffee causally irrelevant to the action. Still, the position is not incoherent (Chalmers 2002, 264), nor is it without defenders (Jackson 2002). This position is compatible with causal closure of the physical, as an epiphenomenon has no causal efficacy. Furthermore, an epiphenomenalist position might be given a naturalistic evolutionary explanation, where the strongly emergent phenomenon can be seen as a byproduct of evolution.

Conclusion

In this paper, I have distinguished between different kinds of emergence and their coherence. I have examined possible examples of strong emergence and discussed the consequences strongly emergent phenomena would have, were they to interact with other phenomena. Finally, I have used a critique of strong emergence to exclude possibilities from philosophy of mind.

I can conclude that strong emergence needs further defense or exemplification in order to become an attainable position and that strong emergentists are left with very few options, the most prominent being epiphenomenalism, in current philosophy of mind.

References

Alexander, Samuel. 1920. *Space, Time, and Deity: The Gifford Lectures at Glasgow 1916-1918*. New York: Dover Publications.

Butterfield, Jeremy. 2011. "Emergence, Reduction and Supervenience: A Varied Landscape." *Foundations of Physics* 41: 920-959.

Chalmers, David J. 2002. "Consciousness and Its Place in Nature." *Philosophy of Mind: Classical and Contemporary Readings*, edited by David J. Chalmers, 247-272. New York: Oxford University Press USA.

Chalmers, David J. 2006. "Strong and Weak Emergence." *The Re-Emergence of Emergence: The Emergentist Hypothesis From Science to Religion*, edited by P. Davies and P. Clayton, 244-254. Oxford: Oxford University Press.

Chalmers, David J. 2010. *The Character of Consciousness*. New York: Oxford University Press.

Ellis, George F. R. 2006. "On the Nature of Emergent Reality." *The Re-Emergence of Emergence: The Emergentist Hypothesis From Science to Religion*, edited by P. Davies and P. Clayton, 79-107. Oxford: Oxford University Press.

Goff, Philip. 2006. "Experiences Don't Sum." *Journal of Consciousness Studies* 13 (10- 11): 53–61.

Seager, William, Philip Goff and Sean Allen-Hermanson. 2020. "Panpsychism." *The Stanford Encyclopedia of Philosophy*. Accessed 1st of June 2020. https://plato.stanford.edu/entries/panpsychism/

Jackson, Frank. 2002. "Epiphenomenal Qualia." *Philosophy of Mind: Classical and Contemporary Readings*, edited by David J. Chalmers, 273-280. New York: Oxford University Press USA.

Kim, Jaegwon. 1999. "Making Sense of Emergence." *Philosophical Studies* 95 (1-2): 3–36.

Kim, Jaegwon. 2011. *Philosophy of Mind*. Boulder: Westview Press. 3. edition.

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

Nagel, Thomas. 1979. *Mortal Questions*. Cambridge: Cambridge University Press.

Ney, Alyssa. 2014. *Metaphysics: An Introduction*. New York: Routledge.

O'Connor, Timothy and Yu Wong, Hong. 2020. "Emergent properties." *The Stanford Encyclopedia of Philosophy*. Accessed 1st of June 2020. https://plato.stanford.edu/entries/properties-emergent/

O'Connor, Timothy and Yu Wong, Hong. 2005. "The Metaphysics of Emergence." *Nous* 39 (4): 658-678.

Strawson, Galen. 2006a. "Realistic Monism – Why Physicalism Entails Panpsychism." *Journal of Consciousness Studies* 13 (10-11): 3–31.

Strawson, Galen. 2006b. "Panpsychism? Reply to Commentators, with a Celebration of Descartes." *Consciousness and its place in nature: does physicalism entail panpsychism?* edited by A. Freeman, 184–280. Imprint Academic

Strawson, Galen. 2016. "Mind and Being: The Primacy of Panpsychism." *Panpsychism: Contemporary Perspectives*, edited by Godehard Brüntrup and Ludwig Jaskolla, 75–112. Oxford: Oxford University Press.