EXPLANATORY EMERGENCE AS A GUIDE TO METAPHYSICAL STRUCTURE

Elanor Taylor

ABSTRACT

The view that there is a distinction between strong, or metaphysical, emergence and weak, or explanatory, emergence, is widely accepted. It is natural, on this view, to regard accounts of strong and of weak emergence as performing different kinds of work, such that strong conceptions of emergence help us to uncover metaphysical structure, while weak conceptions of emergence help us to understand the limits of our scientific explanations. If we accept this division of labor, then it appears that we cannot use an account of weak, explanatory emergence to find out about metaphysical structure.

In this paper, I explore the view that explanatory conceptions of emergence cannot be used for metaphysical purposes, and argue that it is false. Even those who reject strong emergence can, at least in principle, use certain explanatory accounts of emergence as guides to metaphysical structure. On this approach, emergence itself is explanatory, but explanatory emergence may sometimes obtain for metaphysical reasons.
1. Introduction

In the emergence literature it is widely accepted that there is a distinction between strong, or metaphysical, emergence and weak, or explanatory, emergence.¹ Strong emergents are metaphysically autonomous from their bases, where this autonomy is understood differently in different accounts, sometimes in terms of a difference in causal powers, modal status, or fundamentality.² Weak emergents, on the other hand, are not metaphysically autonomous from their bases, but are instead merely explanatorily autonomous from their bases, and the details about autonomy also vary across different accounts. Weak emergents may be impossible to explain in terms of their bases, for example, or they may require an unusual explanatory approach.³ There are other taxonomies of emergence, but this distinction between strong and weak emergence is intuitive and widely accepted.⁴

It is natural to regard strong and weak accounts of emergence as performing different kinds of work. On this view, we use strong conceptions of emergence to uncover metaphysical structure, while we use weak conceptions of emergence to understand the limits of our scientific knowledge and explanations. If we accept this division of labor, then it appears that we cannot use a weak account of emergence to make

---

⁴ An exception to this general acceptance is Carl Gillett, who argues that the distinction between strong and weak emergence is a “false dichotomy”. See discussion in Gillett, C. (2016). Jessica Wilson also recommends understanding both weak and strong emergence in metaphysical terms, in Wilson, J. (2015)
metaphysical claims, because although metaphysicians may recognize and use weak conceptions of emergence, they do not use them to find out about metaphysical structure. Furthermore, some philosophers argue that strong, metaphysical conceptions of emergence are problematic. If they are right, then apparently all we can do with an account of emergence is describe certain features of scientific practice, and the limits of our scientific explanations.

In this paper, I explore the view that explanatory conceptions of emergence cannot be used for metaphysical purposes, and argue that it is false. Even those who reject strong accounts of emergence can, at least in principle, use certain explanatory accounts of emergence as guides to metaphysical structure. On this approach, emergence itself is explanatory, but explanatory emergence may sometimes obtain for metaphysical reasons. As I will argue, showing that certain cases of

---

5 One could argue that we can and do already use weak explanatory conceptions of emergence for metaphysical purposes. For example, we could use them to find out about identities, depending on whether or not identity is explanatory. If identity is explanatory, then the absence of an explanation of the instantiation of some higher-level property in terms of the instantiation of some lower-level property could indicate that the properties are not identical. If identity is not explanatory, then the absence of an explanation of the instantiation of some higher-level property in terms of the instantiation of some lower-level property could follow from the fact that the two properties are identical, rather than that they are metaphysically distinct. I agree that weak explanatory emergence can be used in metaphysical inquiry, but in this article I will show that moving from unavailable explanation to a metaphysical claim is not so straightforward. In order to move from unavailable explanation to metaphysical claim, we must *metaphysically interpret* the unavailable explanation, and work out which absent explanations are metaphysically significant. The rest of this piece is devoted to discussion of how such inquiry could proceed.

6 Kim, J. (2006); Taylor, E. (2015b)
explanatory emergence obtain for metaphysical reasons will be difficult, and will involve challenging questions about the nature of explanation and the extent to which explanation and the absence of explanation is metaphysically significant. But in principle, an explanatory conception of emergence can be used in metaphysical inquiry.

In Section 1 I explore the distinction between strong and weak emergence, and certain challenges to strong emergence. In Section 2 I introduce my own explanatory conception of emergence, and in Section 3 I show that this approach to emergence can be used in metaphysical inquiry, but that doing so requires answers to some difficult questions about the nature of explanation. In Section 4 I discuss two case studies, the 19th century debate between mechanists and vitalists and the contemporary debate about the explanatory gap in philosophy of mind, to illustrate the use of explanatory emergence as a guide to metaphysical structure.

2. Weak and strong emergence

The claim that a given property is emergent is made relative to a distinction between micro- and macro-level properties. For example, some have argued that properties of our conscious experience (macro) emerge from properties of our brains (micro), while others have argued that properties of certain systems (macro) emerge from the properties of those system’s components (micro). The nature of this micro-macro distinction varies across different purported cases of and accounts of emergence. There are many different accounts of emergence, and much diversity in the emergence literature. However across this diversity there is a rough, shared schema for emergence: given a distinction

---

7 I will presume that emergents are properties.
between micro- and macro-level properties, emergent properties are macro-level properties that are both dependent upon and autonomous from certain micro-level properties. This is under-specified, and so part of the challenge of developing an account of emergence is to articulate the precise nature of this dependence and autonomy.

One commonly accepted taxonomy of emergence arises out of two different ways of understanding emergent autonomy. On this picture, there are two different kinds of emergence: metaphysical, or strong, emergence and explanatory, or weak, emergence. Metaphysical emergents are in some way metaphysically autonomous from their bases, (the micro-level properties that give rise to them). For example, metaphysical emergents may have causal powers that their bases do not have, they may not be necessitated by their bases, or they may be dependent upon their bases while also being fundamental. Explanatory emergents are not metaphysically autonomous from their bases, but are instead merely explanatorily autonomous from their bases. For example, explanatory emergents may be difficult, or even impossible, to explain in terms of features of their bases, or may require us to adopt new explanatory frameworks.

Some philosophers have argued that the concept of weak emergence is most commonly appealed to in scientific practice, while the concept of strong emergence is most commonly appealed to in philosophy, and particularly in metaphysics.

There are other taxonomies of emergence. Jessica Wilson, for example, draws a distinction between weak and strong emergence while

---

treated both as metaphysical (2015). Paul Humphreys has argued that there is a third, computational, kind of emergence, and that emergence is diachronic, rather than synchronic (2008, 2016). Furthermore, many reject the idea that scientific practice mostly appeals to the concept of weak emergence, not least because many treatments of emergence in physics take emergence to be metaphysical.\textsuperscript{12} But this distinction between strong and weak emergence, and the corresponding idea that strong emergence belongs in metaphysics, while weak emergence belongs in philosophy of science and scientific practice, is widely accepted.

Some philosophers have argued that accounts of strong emergence are problematic. For example, Jaegwon Kim has argued that strong emergence requires downward causation, but that downward causation is impossible, and so that strong emergence is impossible (2006). I have argued that accounts of strong emergence face the ‘collapse problem’, which shows that there is no non-arbitrary way to distinguish between macro- and micro-level properties in cases of emergence, and that the only way to preserve a non-arbitrary conception of emergence is to understand it in purely explanatory terms (Taylor, 2015b). In addition to these worries, strong emergence is ontologically unparsimonious and apparently mysterious, all of which suggests that we should avoid strong emergentism.

If we endorse the distinction between strong and weak emergence, and also accept these arguments against strong conceptions of emergence, then it may appear that the only thing we can reasonably do with a concept of emergence is explore and articulate certain features of scientific practice. If we do not accept the arguments against the strong conceptions of emergence, but do accept the distinction between weak

\textsuperscript{12} For discussion of ontological emergence in physics, see Morrison, M. (2012) and Silberstein, D. and McGeever, M. (1999)
and strong emergence, then we are still left with the apparent result that we can only practice metaphysics using a strong conception of emergence. On either position, we cannot use an explanatory account of emergence to make claims about metaphysical structure, though metaphysicians may use weak emergence for other purposes. In this paper I will show that this view is false. We can use an explanatory conception of emergence to make metaphysical claims. Even those who are skeptical about strong, metaphysical emergence can do so, because even if emergence itself is entirely explanatory, explanatory emergence may obtain for metaphysical reasons. In the following section I will introduce and discuss my explanatory approach of emergence, and in Sections 3 and 4 I will show how this account can be used in metaphysical inquiry.

3. The explanatory conception of emergence

I have defended the following account of emergence, which I will refer to as ‘the explanatory conception’:

Given components A, B, C... n arranged in relation r into a whole, and an observer O, property x of the whole is emergent for O iff there is no scientific explanation available to O of the fact that the following regularity obtains of natural necessity: Whenever components A, B, C...n are combined in relation r, the resulting whole instantiates property x. (Taylor 2015a, 659)

13 Though see discussion in footnote 4.
According to the explanatory conception emergence is relativized to a number of different factors including an observer, a form of explanation, a standard for unavailability and a distinction between component and whole. What counts as a component, a whole, an explanation, unavailability, and an observer are left open, so each case of emergence is relative to a specification of each of these variables. I leave these variables open rather than specify particular kinds of explanation, observers and standards for unavailability to which emergence is always relative, because doing so permits a more unified account of emergence that can accommodate the diversity evident in scientific and philosophical use of the concept.\textsuperscript{14}

According to the explanatory conception observers can converge on cases of emergence, such that a certain property may be emergent for a group, or even all observers, while other properties may be emergent for only one observer. Similarly, some cases of emergence may obtain relative to only one standard of unavailability, while others may obtain across many, or all, standards of unavailability. Some cases of emergence may obtain only relative to one kind of explanation, while others may obtain across many, or even all, kinds of explanation. Accordingly, on this view there are many different kinds of cases of emergence. Some cases of emergence will be scientifically and philosophically uninteresting, such as those cases relative to odd standards of unavailability such as, “the observer cannot remember the explanation at the moment”, or relative to observers who, for example, lack certain mathematical skills. However, other cases of emergence will be of great scientific and philosophical interest, such as those that obtain for all observers, or for all observers in the contemporary scientific community, for a scientifically significant standard of unavailability and type of explanation.

\textsuperscript{14} Discussed in Taylor, E. (2015a) Sections 6 and 7.
My goal in this discussion is to show that an explanatory conception of emergence can be used for metaphysical purposes, rather than to defend this particular account of emergence. But I will briefly mention two reasons to endorse the explanatory conception: it offers a unified explication of the concept of emergence, and it avoids a serious problem, the ‘collapse problem’.15

Emergence is commonly associated with the unavailability of certain forms of explanation, but the kind of explanation associated with emergence varies across different scientific disciplines, and different areas of philosophy. For instance, philosophers of mind exploring the claim that properties of conscious experience emerge from properties of the brain often focus on certain forms of reductive explanation, particularly a priori reductive explanation.16 Philosophers and scientific practitioners working in Artificial Life have focused on cases of emergence relative to certain forms of derivation without simulation.17 Cases of emergence in biology and in chemistry are often held to be relative to the unavailability of certain kinds of non-systemic explanation.18 Furthermore, the relevant observers vary, as some philosophers are interested in cases of emergence that obtain for idealized observers, such as a perfectly rational observers with complete scientific knowledge, whereas scientific practitioners are often more interested in cases that obtain relative to actual members of the contemporary scientific community. This level of diversity in the use of the concept of emergence has led to pessimism about the prospects for a

15 The first is defended in Taylor, E. (2015a), the second in Taylor, E. (2015b)
16 Chalmers, D. and Jackson, F. (2002); Block, N. & Stalnaker, R. (1999) See also discussion in Section 4.1.
18 Nitschke, J. (2009)
unified explication of emergence, but the explanatory conception offers a unified explication that can accommodate the diversity.\textsuperscript{19}

Accounts of emergence face a serious problem called the ‘collapse problem’, which can be avoided by endorsing the explanatory conception.\textsuperscript{20} For an illustration of the collapse problem, consider the claim that a property of qualitative experience emerges from the properties of some non-conscious brain parts. Imagine a group of neurons arranged in a configuration, which we can call \textit{p}, and stipulate that whenever \textit{p} is instantiated, property of qualitative experience \textit{q} is also instantiated.\textsuperscript{21} Now imagine a philosopher who holds that if \textit{q} is emergent from the properties of the neurons, then from full knowledge of those properties and the fact that the neurons are arranged in \textit{p}, we cannot deduce the fact that \textit{q} is instantiated. She thinks that \textit{q} is emergent, and so that \textit{q} meets this condition. However, the philosopher faces a problem: Each of the neurons has the property of giving rise to qualitative experience with property \textit{q} when gathered with other neurons into configuration \textit{p}. If this is included among the properties of the neurons then it is very simple to deduce that \textit{q} is instantiated, and \textit{q} does not meet the condition for emergence. The emergence has ‘collapsed’ with the inclusion of these properties among the properties of the neurons. This ‘collapse problem’ generalizes and applies to many accounts of emergence. The threat of the collapse problem is that it appears to show that emergence is a relationship between merely arbitrary groups of properties, rather than a genuine distinction in nature. An intuitive and apparently obvious attempt to solve this

\textsuperscript{19} I defend this in Taylor, E. (2015a)
\textsuperscript{20} Taylor, E. (2015b)
\textsuperscript{21} This is obviously a simplified example, not intended to be neuroscientifically accurate.
problem, by stipulating that the micro-level should only include properties from a certain privileged group, also fails.\textsuperscript{22} However, we can avoid the collapse problem and preserve the idea that emergence is a non-arbitrary phenomenon by adopting the explanatory conception.\textsuperscript{23} Accordingly, there are good independent reasons to endorse the explanatory conception.

4. Explanatory emergence as a guide to metaphysical structure

It may appear that the most we can do with the explanatory conception of emergence is document the limits of our explanations, and perhaps also clarify certain aspects of scientific practice. If you agree that the explanatory conception is the only view of emergence that avoids the collapse problem, for instance, then you might also think that simply no metaphysical work can be done with an account of emergence. However, this is not true. In the rest of this paper I will show that we can use the explanatory conception as a guide to metaphysical structure. My verdict will be cautious, because, as I will show, some apparently intuitive and straightforward ways to metaphysically interpret explanatory emergence turn out to fail. It is possible to use the explanatory conception as a guide to metaphysical structure, but this will be difficult, and will require us to engage with deep and challenging questions about the nature of explanation.

Emergence, according to the explanatory conception, is the unavailability of a certain kind of explanation to an observer or

\textsuperscript{22} See Section 3 of Taylor, E. (2015b)
\textsuperscript{23} See Section 4 of Taylor, E. (2015b)
observers. To *metaphysically interpret* emergence is to take the unavailability of the explanation as evidence for, or a reason to endorse, a metaphysical claim. Given micro-level property/ies $a$ and macro-level property $b$, and the unavailability of an explanation of $b$ in terms of $a$, one could argue that the absence of that explanation indicates that we should be dualists about $a$ and $b$, or that $b$ is as fundamental as $a$, or that $b$ and $a$ belong to distinct levels in nature. The history of philosophy contains many attempts to take the unavailability of certain explanations as evidence for a variety of such metaphysical claims. For example, some of the British Emergentists, certain vitalists (who will be discussed in section 5), and some participants in debates about the explanatory gap (which will also be discussed in section 5) have taken the absence of certain explanations as evidence for metaphysical positions such as dualism or vitalism.\(^{24}\) To metaphysically interpret explanatory emergence just is to take the absence of an explanation as evidence for a metaphysical thesis.

On the explanatory conception, many cases of emergence will obtain for merely epistemic reasons. For example, certain cases of emergence will obtain because a particular observer lacks relevant knowledge that would allow them access to an explanation, such as observers with no knowledge of calculus to whom explanations that appeal to mathematical models will be unavailable. There may also be cases of emergence in which an explanation is unavailable to an observer or observers because they lack a particular tool, such as a certain kind of computer. Such cases are clearly merely epistemic, and should not be given a metaphysical interpretation. However, it is possible that certain cases of explanatory emergence do obtain for metaphysical reasons. These may be cases in which the relevant observers are metaphysically

privileged in some way, or perhaps cases that obtain for all possible observers. They may be cases of emergence relative to a metaphysically privileged notion of explanation, or perhaps for all explanations, or cases of emergence relative to a metaphysically significant standard of unavailability. In the rest of this section I will describe the questions about explanation, observers and unavailability that must be answered in order to metaphysically interpret certain cases of emergence, suggest some answers to such questions and reject some others. I will not attempt to settle the question of which cases of explanatory emergence are metaphysically significant, but will point towards some particularly promising lines of inquiry.

4.1 Explanation

According to the explanatory conception of emergence, emergence involves the unavailability of certain kinds of explanation. In this section I will discuss some different types of explanation, and the kinds of questions we must answer about explanation in order to metaphysically interpret certain cases of explanatory emergence.

I will make two presumptions. The first is that neither explanatory realism nor explanatory anti-realism is true. Explanatory realism is the position that all explanations succeed by giving information about whatever metaphysically determined the explanandum, where metaphysical determination is a kind of metaphysical production or responsibility, of which causation and grounding are examples. Explanatory anti-realism is the position that no explanation succeeds by giving information about whatever metaphysically determined the

---

25 For discussion see Audi, P. (2012); Audi, P. (2015); Kim, J. (1988); Ruben, D-H. (1990)
explanandum, because explanation succeeds in virtue of epistemic factors, such as generating understanding in the person who asks for the explanation.26 There are good independent reasons to reject both of these positions. As I have argued elsewhere, even the most apparently plausible versions of explanatory realism face serious counterexamples, and the explanatory realist fails to accommodate the pragmatic and diverse nature of explanation (Taylor, 2017). Explanatory anti-realism also faces counterexamples, and fails to make sense of the role that explanation plays in metaphysics, including the central role of inference to the best explanation in metaphysics. Furthermore, there are good dialectical reasons to reject both of these positions. If explanatory anti-realism were true, then the availability and unavailability of explanations would be metaphysically insignificant, as no explanation would give information about metaphysical structure. So if explanatory anti-realism were true, it would be impossible to us explanatory emergence as a guide to metaphysical structure. If explanatory realism were true, then the availability and unavailability of any kind of explanation would be metaphysically significant, because all explanations would give information about metaphysical structure. So, on this view metaphysically interpreting explanatory emergence would be fairly straightforward (though there would still be interesting questions about unavailability and the reasons for the unavailability of certain explanations, to be discussed in the next section).

The second presumption is that the distinction between the explanations that give information about metaphysical structure and those that do not is not straightforward, and so that it is not always easy to tell which group a certain explanation falls into. One good reason to adopt this view is that it is reflected in recent debates about noncausal

---

26 For discussion see Van Fraassen, B. (1980); Achinstein, P. (1983)
explanation and grounding explanation. There is a substantial literature on whether certain explanations are causal or noncausal, and a similarly substantial literature about whether there are grounding explanations. Given that causal explanations and grounding explanations are kinds of explanation that give information about metaphysical structure (causal structure and grounding structure respectively) the mere existence of such debates indicates that working out whether or not a given explanation is metaphysically significant is not straightforward. So we should not expect it to be easy to work out which absent explanations are metaphysically significant, and hence which cases of explanatory emergence obtain for metaphysical reasons. Accordingly, the view I will adopt is a form of pluralism, according to which some explanations give information about metaphysical structure, and some explanations do not, and where the distinction between these two kinds of explanation is not always obvious.

The question at hand is: what kinds of explanation, if any, can tell us about metaphysical structure? More precisely: are there certain kinds of explanation the unavailability of which can be interpreted as a case of explanatory emergence that obtains for metaphysical reasons?

In exploring which forms of absent explanation, and hence which versions of explanatory emergence, obtain for metaphysical reasons, one apparently appealing strategy is to argue that emergence obtains for metaphysical reasons only in cases where the emergence obtains relative to every form of explanation. This is a natural way to metaphysically interpret emergence, as in taking this approach we avoid having to specify which kinds of explanation are and are not metaphysically significant. It is also intuitively plausible that the absence of any kind of explanation of a certain macro-level property in terms of certain micro-level properties would indicate a metaphysical difference between the

---

two. However, this strategy faces problems. In most cases of emergence some kind of explanation is unavailable, but not all kinds of explanation. For example, even those who argue that phenomena such as consciousness or free will are metaphysically emergent also often hold that we can incorporate such phenomena into our understanding of the world and hence explain them non-reductively by positing psychophysical laws, or alternative nonreductive forms of explanation.\textsuperscript{28} In scientific practice, emergent phenomena are not typically thought to be unexplained for all standards of explanation, merely unexplained for some particularly scientifically significant standard for explanation.\textsuperscript{29}

The philosopher who adopts this apparently intuitive strategy of focusing on only those cases of emergence that obtain for all kinds of explanation has to make sense of this practice.

Some philosophers have argued that certain kinds of explanation that involve an a priori conceptual connection between the explanans and explanandum are privileged guides to metaphysical structure. According to proponents of the view that there is an explanatory gap between the facts about conscious experience and physical facts about the brain, the absence of a certain kind of conceptual explanation of facts about conscious experience in terms of facts about the brain is metaphysically significant, and indicates that consciousness is metaphysically distinct from brain activity. Because we cannot explain mental states and concepts in physical terms, these authors suggest, physicalism must be false. For example, Frank Jackson and David Chalmers (2002) argue that a priori reductive explanation is metaphysically significant, where a priori reductive explanation is a kind of explanation in which certain

\textsuperscript{28} A clear example of this is in Chalmers, D. (1996) Some British Emergentists also endorsed the view that there are emergent laws. See McLaughlin, B. (1992)

\textsuperscript{29} Such as Morrison, who argues that physical phenomena such as superconductivity are emergent on these grounds. See Morrison, M. (2012)
facts are reduced to other facts through an a priori conceptual connection. According to Jackson and Chalmers, the absence of an a priori reductive explanation of any fact in terms of physical facts indicates that physicalism is false, and the facts about consciousness are not so explained by the physical facts. This claim generated a debate about the correct definition of physicalism, with some commentators arguing that Jackson and Chalmers’ definition of physicalism is simply too strong. Ned Block and Robert Stalnaker (1999) argue that physicalism need not require that all facts be a priori entailed by physical facts, and so that the absence of an a priori reductive explanation of facts about consciousness in terms of physical facts does not indicate that physicalism is false. They argue that many successful scientific reductive explanations do not involve a priori analysis, and so that there is no reason to expect a reductive explanation of consciousness to meet this standard. Jackson and Chalmers (2002) responded by precisifying the role of a priori analysis in their conception of reductive explanation, and by challenging Block & Stalnaker’s case studies, particularly the case of life. This continued into a lengthy debate about the correct definition of physicalism, which continues today.

Perhaps, then, we should take a priori reductive explanation as a guide to metaphysical structure, and along with Jackson and Chalmers take the unavailability of a priori reductive explanations as metaphysically significant. However, we do not need to explore the details of the debate about a priori reductive explanation and the formulation of physicalism to recognize that there are good reasons to be sceptical a priori physicalism, and accordingly to be sceptical about the metaphysical implications of a priori reductive explanation. The standard that physicalism must involve an a priori derivation of all facts from the physical facts is unfeasibly strong, not least because it is not obvious that such an a priori reductive explanation obtains between
physical facts, let alone between physical and non-physical facts.\textsuperscript{30} There is much debate in philosophy of physics about the relationship between statistical mechanics and thermodynamics, and about whether thermodynamics reduces to statistical mechanics.\textsuperscript{31} It is non-naturalistic to define physicalism as involving a kind of a priori connection between physical and non-physical facts that does not clearly obtain between physical facts themselves. Furthermore, as I have argued elsewhere, the view that deductive explanations are metaphysically significant is poorly motivated, especially in discussions of consciousness and the explanatory gap (Taylor, 2016). Such considerations give us good reason to be skeptical about giving a priori reductive explanation such a central role in the definition of physicalism, and hence as a privileged guide to metaphysical structure.

Another form of explanation that some philosophers argue is metaphysically significant is \textit{grounding explanation}. There are many different accounts of grounding, but the basic idea uniting them is that grounding is a non-causal form of metaphysical determination that obtains between facts, which is either explanatory or bears some kind of close relationship to explanation, such as by supporting explanations.\textsuperscript{32}

\textsuperscript{30} This depends on the definition of “physical facts”, but Jackson and Chalmers have no principled basis in their view for excluding facts about thermodynamics from the set of physical facts.

\textsuperscript{31} See discussion in Sklar, L. (1993) and Callender, C. (1999). In Sklar, L. (2015), Sklar warns that “the complexity of the inter-relationship between the theories should make the philosopher cautious in using this relationship as a well understood and simple paradigm of inter-theoretic reduction.”

\textsuperscript{32} See Correia, F. and Schneider, B. (eds) (2012) for a collection of recent work on grounding. The variety of different accounts of grounding is important, as there are substantial differences between the different accounts with respect to the nature of grounding, the relata of grounding, and the uses to which grounding
Purported examples of grounding include the relationship between facts about categoricals and facts about dispositions (e.g. the fact that a vase is fragile is grounded in facts about the vase’s categorical physical properties), and the relationship between facts about determinates and facts about determinables (e.g. the fact that a certain plant is red is grounded in the fact that it is crimson). In recent work on grounding, Paul Audi (2012) has argued that the availability of certain kinds of non-causal explanation indicates that there is metaphysical grounding. According to Audi, there are non-causal explanations (such as the explanation of facts about determinables by facts about determinates), and all explanation is supported by instances of metaphysical determination. So there must be a non-causal form of metaphysical determination, and that is grounding (Audi 2012, 688). In this argument, Audi takes the availability of certain kinds of non-causal explanation as a guide to metaphysical structure.

Many philosophers have criticized accounts of metaphysical grounding. For example, Daly (2012) has argued that grounding is unintelligible, while Wilson (2014) has argued that grounding is superfluous, as the work proposed for grounding is already covered by familiar metaphysical notions such as realization. However, even if there are viable accounts of grounding, there are reasons to be skeptical about taking grounding explanation as a privileged guide to metaphysical structure, and in particular interpreting the absence of grounding explanations as examples of emergence that obtains for metaphysical reasons. First, to interpret the availability of non-causal explanations as indicating that there is a form of non-causal metaphysical determination, as Audi does, requires us to accept that all explanation is supported by some form of metaphysical determination, either causal or non-causal.

---

can be put. Ney, A. (2016) discusses the impact of these differences with respect to the application of grounding to debates in the philosophy of mind.
have argued elsewhere that this position, a kind of explanatory realism, is false, and that it fails to do justice to a wide variety of actual explanatory practices (Taylor, 2017). Second, grounding explanation is not well suited to a central role in the metaphysical interpretation of explanatory emergence. When metaphysically interpreting explanatory emergence, we are taking the absence of a certain kind of explanation as evidence of a claim about metaphysical structure. However, the absence of a grounding explanation between two facts need not mean that there is a robust and interesting metaphysical distinction between these two facts. There may be no grounding relationship, but there may be a causal relationship, or an identity. So interpreting the unavailability of grounding explanations as a case of explanatory emergence that does pick out metaphysical structure would not be straightforward.

These are examples of attempts to tie explanation to metaphysical structure, and to argue that the presence or absence of certain kinds of explanation can be used as a guide to that structure. In order to use explanatory emergence as a guide to metaphysical structure, then, we must work out which, if any, kinds of explanation are metaphysically significant. In this section I have discussed two attempts to show that certain kinds of explanation are a privileged guide to metaphysical structure, and so that the absence of such an explanation can be interpreted as a case of explanatory emergence that obtains for metaphysical reasons. These attempts faced problems, and more research into the nature of explanation is needed before we can establish which forms of explanation are metaphysically significant.

4.2 Unavailability

In order to use the explanatory conception of emergence in metaphysical inquiry, we also need to establish whether some particular standard of unavailability is metaphysically privileged.
Cases of emergence relative to certain standards for unavailability clearly should not be interpreted as metaphysical. For example, there are cases in which an explanation is unavailable to an observer because that observer lacks a certain tool or specialist knowledge that is available to other observers, such as kind of computer, or a theoretical framework. Because the explanation is available to other observers the emergence should not be interpreted as obtaining for metaphysical reasons. We can automatically rule out any standard of unavailability according to which the explanation is available to other observers at the same time. Similarly, we can rule out any case in which the unavailability is fragile, in that it can quickly be made to disappear with a small discovery or the observer learning a new skill. Such cases are too contingent and relative to the situation of particular individuals to be useful indicators of features of metaphysical structure. Any case in which the standard for unavailability is problematically unscientific should also be ruled out, such as when an explanation is “unavailable by crystal-scrying”, or “unavailable from an astrological perspective”, because such standards for unavailability do not reflect the full range of scientific explanations that may be available relative to different standards.

One intuitive desideratum for the unavailability standard when using explanatory emergence for metaphysical purposes is that the standard of unavailability should be unavailability in principle rather than merely in practice. If an explanation is unavailable in principle, then no further inquiry could help us to develop the explanation, whereas if the explanation is unavailable in practice, then it may eventually become available. If we adopt the standard that emergence obtains for metaphysical reasons only in cases relative to an in-principle

---

35 On a relativistic approach to metaphysics this would look different, but I am presuming a standard, non-relativistic conception of metaphysics according to which metaphysical entities are not context- or observer-dependent.
unavailability standard, then we would avoid the fragile, contingent cases of emergence described in the paragraph above. For example, consider a case of emergence in which a scientific explanation of some phenomenon is available in principle in that some fact about the world explains the phenomenon, but no observer has yet uncovered the explaining fact, and so the explanation is not available in practice to any current observer. This case may be scientifically interesting, but once our scientific knowledge develops, the explanation will be uncovered and emergence relative to the in-practice standard of unavailability will disappear. Accordingly, it may seem sensible to only focus on those explanations that are unavailable in principle. However, doing so generates a significant epistemological problem about what would count as evidence for the unavailability of an explanation in principle, and so merely stipulating that the explanation should be unavailable in principle does not in itself provide a simple route to a metaphysical interpretation of emergence.

4.3 Observers

According to the explanatory conception, observers can converge upon cases of emergence. Certain cases of emergence may obtain for only one or a few observers, others may obtain for entire scientific communities, while others may obtain for all possible observers. This generates a wide range of different kinds of cases of emergence, relative to different kinds of observers. If we want to metaphysically interpret explanatory emergence, then we need to work out which observers are metaphysically significant, and so which cases of explanatory emergence obtain for metaphysical reasons.

There are many cases of emergence that cannot be used as an indication of metaphysical structure because they are relative to metaphysically non-significant observers. For example, if an explanation is unavailable to an observer because that observer contingently lacks
access to a certain tool or a theoretical framework that would make the explanation available, then we should not interpret this as a metaphysically significant case of emergence, because it will disappear when the observer gains access to the relevant explanation. Such cases are simply too fragile to tell us much about the structure of the world. In the section on unavailability I discussed cases of emergence in which explanations are unavailable to particular individuals, and recommended against metaphysically interpreting these.

Certain observers will be non-significant for the purposes of metaphysically interpreting emergence. For example, cases of emergence that are relative only to children, or to people without access to scientific institutions, should not be given a metaphysical interpretation as the unavailability of the relevant explanations in these cases of emergence can be explained in terms of the observer’s lack of education and resources. Similarly, cases of emergence relative to historic observers, such as pre-Daltonian chemists, should also not be interpreted metaphysically as the absent explanation in such cases is explained by the incomplete scientific knowledge of the observers. Such cases are fairly straightforward.

However, interesting questions come up about how closely the scientifically significant and metaphysically significant observers may differ or coincide. For a chemist, the salient and scientifically significant observers are likely to be the members of the contemporary community of practicing chemists, or depending on the case, members of the broader contemporary scientific community. For certain scientific purposes, salient observers may be restricted to members of a particular lab, or observers with access to a particular scientific tool. When thinking about the limits of scientific explanation more generally, the salient observers will be members of the contemporary scientific community in general. However, when using the explanatory conception of emergence in the practice of metaphysics, we must answer a difficult further question: are
the observers that are metaphysically significant the same as the observers that are scientifically significant, or are they a different group?

Answering this question will involve exploring the nature of metaphysics and the relationship between metaphysics and science. On a highly naturalistic view of metaphysics it would be reasonable to think that the metaphysically privileged observers simply are the scientifically privileged observers.\(^{34}\) If metaphysics is in some sense autonomous from scientific practice, then the metaphysically significant observers may be different from the scientifically significant observers.\(^{35}\) I will not attempt to settle these questions here, but I will suggest that the view that metaphysics is completely autonomous from science is problematically non-naturalistic, and so that there must be some overlap between those observers that are scientifically significant and those observers that are metaphysically significant. However, there are equally good reasons to think that the metaphysically and scientifically significant observers will not completely overlap. Much scientific inquiry has non-metaphysical goals, such as creating improved materials and curing diseases. The extent to which success in such goals is dependent upon correctly identifying metaphysical structure is a key question in debates about scientific realism, but it is reasonable to think that because of this pragmatic orientation not all scientific observers will be significant for the purposes of interpreting explanatory emergence. However, some scientific practice does have metaphysical goals, and the scientific observers engaged in such practices will be metaphysically significant.

Another reason to suspect that the scientific and metaphysically significant observers may not entirely coincide is that the metaphysically significant

\(^{34}\) For an articulation and defence of one version of “scientistic metaphysics” see Ladyman, J. and Ross, D. (2009)

\(^{35}\) For a defence of the view that metaphysics is continuous with but autonomous from science, see Paul, L. (2012)
significant observers may be idealized in some way, such as having perfect scientific knowledge, or access to unlimited computational power. Some metaphysicians may be interested in cases of emergence relative to an observer who is a Laplacean demon, for example. These are idealized, hypothetical observers rather than real human beings, and so may be of limited scientific interest even if they are of metaphysical interest. However, basing the metaphysical interpretation of emergence on these hypothetical idealized observers raises some difficult epistemological questions, not least of which is how we can know what explanations are and are not available to a hypothetical idealized observer such as a Laplacean demon or someone with perfect scientific knowledge, whose knowledge and resources are so significantly different from our own.

As in the case of explanation and of unavailability, another way to metaphysically interpret explanatory emergence is to only examine cases of emergence that obtain for all possible observers. Showing that this is the case will, as before, be difficult, but this is another intuitive way to identify the metaphysically significant cases without having to specify a metaphysically significant set of observers. These questions about the metaphysically significant observers are closely connected to questions about the metaphysically significant standards for unavailability and forms of explanation. For example, an idealized observer with complete knowledge will have access to all in-principle available explanations. Accordingly, these questions about the appropriate standard for unavailability, observers and explanations for metaphysical inquiry cannot be addressed entirely independently.
5. Case studies

So far we have seen that it is possible for explanatory emergence to obtain for metaphysical reasons, but establishing that this is the case requires answers to difficult questions about precisely what cases of emergence, involving what observers, explanations and standards for unavailability, can be taken as evidence for any metaphysical claim. In this section I will examine two case studies in which philosophers and scientific practitioners have attempted to use explanatory emergence as a guide to metaphysical structure, though the original authors did not think of their inquiry in terms of emergence. The first case is the 19th century debate between vitalists and mechanists, and the second is the contemporary debate about the explanatory gap in philosophy of mind.

5.1 The vitalist-mechanist debate

The debate between vitalists and mechanists is hard to pin down, both chronologically and philosophically, but for the sake of convenience I will stipulate that I am interested in the 19th century incarnation of the debate. I do not intend to offer a definitive historical interpretation of this debate, but will offer one interpretation of it as a debate about explanatory emergence.

The distinction between the mechanist and the vitalist can be drawn in terms of their respective approaches to the question of life. Beckermann offers this characterization of the distinction:

... mechanisms claimed that the properties characteristic of living organisms (metabolism, perception, goal-directed behavior, procreation, morphogenesis) could be explained mechanistically, in the way the behavior of a clock can be explained by the properties and the arrangement of its cogs, springs and weights.

... vitalists, on the other hand, maintained that the explanation
envisaged by the mechanists was impossible and that one had to postulate a special nonphysical substance in order to explain life—an entelechy or élan vital.

There were different varieties of vitalism and of mechanism, and some have argued that it would be better to think of mechanism and vitalism as collections of strands of thought rather than unified philosophical positions. Benton argues that, “it is a mistake to regard vitalism as one doctrine, or set of doctrines” (1974, 17). For this discussion, however, I will continue to use Beckermann’s definition. C.D. Broad originally presented his account of emergentism as providing middle ground between mechanism and vitalism (1925, 61). This may make it historically confusing to think of the mechanist-vitalist debate in terms of the explanatory conception of emergence, but I will suggest that in this section we put aside this aspect of the history of the term ‘emergence’ and adopt the explanatory conception, so that we can use this case study as an illustration of metaphysically interpreting explanatory emergence.

We can understand this case as a debate about explanatory emergence, and an attempt on the part of the vitalist to metaphysically interpret explanatory emergence. To see how this works, we can begin by applying the explanatory conception to vitalism. Vitalism amounts to the position that for some feature l, a property characteristic of living creatures, and some observer/s, there is no explanation available to the observer/s of the fact that the following regularity obtains of natural necessity: Whenever components $C_1, C_2, \ldots, C_n$ are combined in relation r, the resulting whole instantiates property l. Components $C_1 \ldots C_n$ can be understood as chemical components, but depending on the relevant kind of mechanism, they may alternatively be understood as physical components. In this case, the relevant observers were the members of the contemporary scientific community. For vitalists, the relevant kind of explanation was mechanistic explanation, and they took the absence of mechanistic explanation of certain characteristic properties of living
creatures, such as metabolism and goal-directed behavior, as indicating something about the metaphysical structure of the world. According to the vitalist, given that there was no mechanistic explanation of these features, some other feature of the world must explain them, which, they posited, was the substance élan vital. This is a good example of interpreting explanatory emergence metaphysically—taking the absence of a particular kind of explanation, in this case mechanistic explanation, as a guide to metaphysical structure.

To examine this debate more closely, we can look specifically at the case of fermentation. Louis Pasteur and Marcelin Berthelot offered rival accounts of fermentation, with Pasteur arguing that fermentation was a vital process and Berthelot arguing that fermentation could be explained mechanistically, without needing to call upon the notion of life or of élan vital. Fruton quotes Berthelot stating his objective as follows: “To banish life from all explanations relative to organic chemistry, that is the aim of our studies” (2006, 61) and Berthelot claimed to have developed an entirely chemical (in our sense, mechanistic) explanation of fermentation. Pasteur and Berthelot shared a notion of explanation (mechanistic), of availability (in practice to the contemporary scientific community), and of the relevant observers (the contemporary scientific community), but they disagreed about whether or not an explanation was available according to their shared standards, and so disagreed about whether or not fermentation was emergent (on the explanatory conception of emergence) relative to these shared standards.

This case illustrates the way in which explanatory emergence can be interpreted metaphysically. The vitalist took the apparent absence of a mechanistic explanation of features such as fermentation as indicating that some other fact about the world, the substance élan vital, must explain those features.
5.2 The explanatory gap

The 'explanatory gap' is a label used in philosophy of mind for the idea that we lack a satisfactory explanation for the fact that conscious experience arises from brain activity. Levine offers an intuitive statement of the idea of an explanatory gap:

While we seem to have some idea how physical objects, or systems, obeying physical laws, could instantiate rational and intentional properties, we have no idea, I contend, how a physical object could constitute a subject of experience, enjoying, not merely instantiating, states with all sorts of qualitative character... There seems to be no discernible connection between the physical description and the mental one, and thus no explanation of the latter in terms of the former. (Levine 2001, 76).

The explanatory gap plays a central role in philosophy of mind because philosophers have argued that it generates a serious problem for physicalism. If physicalism is true, then a physicalistically acceptable explanation of conscious experience should be available, and if there is no such explanation, then physicalism must be false. Responses to the challenge to physicalism vary. Some argue that there is an explanatory gap, and that it is a problem for physicalism.36 Some argue that there is an explanatory gap, but that it is not a problem for physicalism.37 Others argue that the explanatory gap does not pose a problem for physicalism, because we can explain the gap through appeal to the idea that phenomenal experience requires particular conceptual resources.38

---

37 Block, N. & Stalnaker, R. (1999)
38 Known as the Phenomenal Concepts Strategy. See e.g. Balog, K. (2012)
As with the case of the vitalist-mechanist debate, we can understand the case of the explanatory gap as an example of an attempt to metaphysically interpret explanatory emergence. The first step is to apply the explanatory conception of emergence to the thesis that there is an explanatory gap. According to the explanatory conception, the view that there is an explanatory gap amounts to the view that for some observer/s and some property $p$ of conscious experience, there is no explanation available to that observer/s of the fact that the following regularity obtains of natural necessity: Whenever components $N_1, N_2, ..., N_n$ are combined in relation $r$, the resulting whole instantiates property $p$, where components $N_1...N_n$ can be understood as neurophysiological components.\(^{39}\)

Debates about the explanatory gap are then taken up with questions about what kind of explanation is appropriate to look for in this case, what the relevant standard of unavailability should be, and who the relevant observers are. Some of the most significant work on this subject, including the dialogue between Jackson & Chalmers and Block & Stalnaker discussed in Section 4.1, focuses specifically on questions about explanation, and whether in order to close the explanatory gap an explanation must involve a priori derivation.\(^{40}\) Those who argue that the explanatory gap is metaphysically significant, in so far as it indicates that physicalism is false, that dualism is true, or that nonreductive physicalism is true, and so on, have used explanatory emergence as a guide to metaphysical structure.

\(^{39}\) Though they could be physical, chemical etc components instead.

6. Conclusion

It is commonly accepted that there is a distinction between strong, metaphysical emergence and weak, explanatory emergence, and in light of this distinction as well as arguments against strong conceptions of emergence, it is tempting to think that a weak, explanatory conception of emergence cannot be used to make metaphysical claims. In this paper I have shown that this is false, by introducing an explanatory conception of emergence and showing how it can be used as a guide to metaphysical structure. Even if emergence itself is explanatory, explanatory emergence may obtain for metaphysical reasons. Showing that this is the case, however, will require answers to challenging questions about the nature of explanation and about what explanation can tell us about the world. 41

Department of Philosophy
John Hopkins University, Baltimore, USA
Email: elanort@jhu.edu

REFERENCES


41 Thanks to the Boise State University Department of Philosophy for helpful discussion, and to an anonymous referee.


Humphreys, P., 2016 *Emergence*. Oxford University Press.


Nagel, T., 1974, “What is it like to be a bat?” *The Philosophical Review*. 83 (4): 435-450


Taylor, E., 2015b, “Collapsing Emergence.” *Philosophical Quarterly* 65 (261): 732-753