Argumenta 10, 1 (2024)
Book Symposium

On Jessica Wilson's *Metaphysical Emergence* OUP 2021

The Journal of the Italian Society for Analytic Philosophy

Précis of *Metaphysical Emergence*

Jessica Wilson

University of Toronto

Introduction

The notion of metaphysical emergence is inspired by certain target cases, whereby—on the face of it, and in ways I'll expand on shortly—'higher-level' entities (objects, events, and the like) and features (properties, relations, behaviours, and the like) cotemporally materially depend on 'lower-level,' ultimately fundamental physical, micro-configurations and features; yet are also to some extent autonomous, ontologically and causally, from dependence base configurations and features. Relatedly, metaphysical emergence is inspired by a conception of natural and artifactual reality as manifesting a kind of leveled structure generally mirrored in the special sciences vis-à-vis the more fundamental physical sciences.

But what is metaphysical emergence, more precisely, and is there more than one variety of such emergence? And is there (really) any metaphysical emergence, in principle and moreover in fact?

In *Metaphysical Emergence* (2021), I aim to provide clear and systematic answers to these questions. I argue that there are two, and only two, forms of metaphysical emergence capable of accommodating the target cases—one 'Weak' (compatible with a physicalist world-view, given that the lower-level goings-on are physical), one 'Strong' (not so compatible). After defending the in-principle viability of each form of emergence, I consider whether complex systems, ordinary objects, consciousness, and free will are actually metaphysically emergent. I argue that some cases of each phenomenon are plausibly Weakly emergent, and I offer a new argument for there being free will of a Strongly emergent variety.

In what follows, I expand upon this rough overview, summarizing each chapter of *Metaphysical Emergence*. In the interest of efficiency, the presentation sometimes mixes prose with features more characteristic of a visually structured outline.[1]

Chapter 1: Key Issues and Questions

In Chapter 1, I begin by canvassing the prima facie motivations for thinking that there is metaphysical emergence (§1.1). To start, scientific orthodoxy takes for

[1] Please keep in mind that this précis necessarily elides what I take to be important dialectical qualifications and content. The book remains the official statement of my view(s).

Argumenta 10, 1 (2024): 191—224
ISSN 2465-2334

DOI 10.14275/2465-2334/202419.wil
First published: 30 November 2024

granted *Physical monism*, understood as contrasting with substance pluralist views such as Cartesian dualism or vitalism:

- *Physical monism*: The only matter or substance is physical matter or substance, such that the matter of a macro-entity at a time is inherited from some micro-configuration of ultimately physical constituents at that time.

Scientific orthodoxy also takes for granted that the features of macro-entities do not float entirely free of features of micro-configurations:

- *Cotemporal dependence*: The features of any macro-entity at a time or over a given temporal interval are at least in part a function of the features of the micro-configuration(s) which materially constitute the macro-entity at that time or during that temporal interval.

Reflecting these commitments, we can say that on the face of it, macro-entities and features *cotemporally materially depend* on micro-configurations and features.

What about autonomy? That macro-entities and features are to some extent both ontologically and causally autonomous from—that is, distinct from and distinctively efficacious as compared to—their underlying micro-configurations and features is motivated by a variety of considerations, including:

- *Distinctive taxonomies*: Special-science entities/features are classified under types which appear to be different from those classifying micro-configurations and features of such configurations (supports distinctness).
- *Distinctive causal laws*: Special-science entities enter into special-science laws describing features and behaviours of, including causal interactions involving, such entities—laws that, on the face of it, are different from those governing physical micro-configurations (supports distinctive efficacy, hence also distinctness).
- *Universal properties and behaviour*: Many special-science entities/features, including thermodynamic complex systems and features, are functionally and causally independent of underlying micro-configurations and features (supports distinctive efficacy, hence also distinctness).
- *Perceptual unity*: Macro-entities such as trees and tables perceptually appear to us as comparatively stable, unified entities, even though (as science tells us) they are materially constituted by complex, constantly changing micro-configurations (supports distinctness).
- *Compositional flexibility*: The existence and persistence of macro-entities/features typically appears to transcend that of underlying micro-configurations, in not depending on any *specific* micro-configuration(s) or features (supports distinctness).
- *Seemingly free will*: It introspectively seems as if we human persons are able to make free choices to produce (or intend to produce) certain effects, where this efficacy appears to be quite different from that associated with the (deterministically or indeterministically) lawfully governed micro-configurations and features upon which we and our mental states cotemporally materially depend (supports distinctive efficacy, hence also distinctness).

On the face of it, then, many macro-entities are *ontologically and causally autonomous* from—that is, distinct from and distinctively efficacious as compared to—

the micro-configurations and features upon which they cotemporally materially depend.

There is thus clear good reason to explore the notion of metaphysical emergence, understood as coupling *cotemporal material dependence* with *ontological and causal autonomy*.

Two key questions are immediately salient (§1.2):

1. Just what is metaphysical emergence, more precisely? How is it, exactly, that macro-entities and features can cotemporally materially depend on micro-configurations and features, while retaining some degree of ontological and causal autonomy? And is there more than one way in which this can be—is there more than one form of metaphysical emergence?
2. Is there actually any metaphysical emergence? To start: are there any insuperable problems with the notion(s) of metaphysical emergence, such that emergence is, at best, an epistemic or representational phenomenon? And supposing that a given variety of metaphysical emergence is in-principle viable, are there any actual cases of such emergence?

Indeed, in past decades there has been an explosion of philosophical and scientific interest in metaphysical emergence; yet the answers to the key questions have remained unclear. In re the first question: a bewildering variety of accounts of metaphysical emergence has been proposed, appealing to different, often incompatible interpretations of the core notions of dependence[2] and autonomy.[3]

[2] Candidate accounts of the dependence at issue in metaphysical emergence include mereological ('part-whole') determination (see Stephan 2002, Gillett 2002), causation or nomological connection (see Searle 1992, O'Connor and Wong 2005), functional realization (see Putnam 1967, Boyd 1980, Poland 1994, Antony and Levine 1997, Melnyk 2003), constitutive mechanism (see Craver 2001, Haug 2010, Gillett 2016), the determinable-determinate relation (see MacDonald and MacDonald 1986, Yablo 1992, Ehring 1996, Wilson 2009), inheritance of causal powers (see Kim 1992, Wilson 1999 and 2015, Shoemaker 2000/2001), and primitive 'Grounding' (see Schaffer 2009, Dasgupta 2014).

[3] Candidate accounts of the ontological and/or causal autonomy at issue in metaphysical emergence include nomological but not metaphysical supervenience (see Cleve 1990, Chalmers 1999, Seager 1999/2016, Noordhof 2010), non-fundamental novelty (of features, powers, laws, entities) (see Humphreys 1996, Wimsatt 1996, Crane 2001, Pereboom 2002, Megill 2013), fundamental novelty (of features, powers, forces/interactions, laws, entities) (see Mill 1843/1973, Alexander 1920, Broad 1925, Kim 1992, O'Connor 1994, Cunningham 2001, Wilson 2002 and 2015, Barnes 2012, Paolini Paoletti 2017), non-additivity/non-linearity (see again Mill, Alexander, and Broad, Newman 1996, Bedau 1997, Silberstein and McGeever 1999, Mitchell 2012), 'downward' causal efficacy (see Morgan 1923, Sperry 1986, Klee 1984, Thompson and Varela 2001, Searle 1992, Schroder 1998, Stephan 2002), multiple realizability/universality/compositional plasticity (see Putnam 1967, Fodor 1974, Boyd 1980, Klee 1984, LePore and Loewer 1989, Wimsatt 1996, Antony and Levine 1997, Aizawa and Gillett 2009, Morrison 2012), causal proportionality/difference-making/counterfactual considerations (see Yablo 1992, LePore and Loewer 1987 and 1989, Bennett 2003), elimination in degrees of freedom (see Wilson 2010 and Lamb 2015), sometimes associated with symmetry breaking (see Morrison 2012), and the holding of a proper subset relation between token powers (see Wilson 1999), sometimes cashed in terms of a proper parthood relation between properties and behaviours (see Shoemaker 2000/2001, Clapp 2001, Rueger and McGivern 2010). Also relevant here are 'epistemic criteria' accounts of ontological and/or causal autonomy, including in-principle failure of deducibility/predictability/explicability (see Broad 1925, Hempel and Oppenheim 1948, Klee 1984, LePore and Loewer 1989), pre-

Indeed, the extent of variability has led many to conclude that there is nothing systematic to be said or discovered about metaphysical emergence. The answer to the second key question has also remained unclear, owing to still-live concerns about whether the appearances of metaphysical emergence are genuine. Among these concerns are that metaphysical emergence is naturalistically unacceptable; that considerations of parsimony push against taking the appearances of metaphysical emergence ontologically seriously; that the notion of metaphysical emergence is either trivially fulfilled or trivially never fulfilled; and—perhaps most problematically—that metaphysically emergent entities or features, were they to exist, would give rise to problematic causal overdetermination of effects already produced by micro-configurations/features. Here the diversity of accounts of emergence again muddies the waters; for while some accounts have resources to respond to some concerns, the absence of any systematic treatment of metaphysical emergence renders it unclear whether the notion can survive all the various attacks.

In light of all this, the point and purpose of my book is to provide clear, compelling, and systematic answers to the two key questions of what, more precisely, metaphysical emergence is, and whether there actually is any such emergence. As discussed in §1.3, I go about this project as follows:

- In Ch. 2, I argue that there are two (and only two) schematic forms of metaphysical emergence which accommodate the target cases. One—'Weak emergence'—is compatible with physicalism, the view that all broadly scientific goings-on are completely metaphysically dependent on lower-level physical goings-on, on the assumption that the lower-level (ultimately compositionally basic) goings-on are physical; the other—'Strong emergence'—is incompatible with physicalism, on that assumption.[4]
- In Ch. 3, I consider and respond to a range of objections to the viability of Weak emergence.
- In Ch. 4, I consider and respond to a range of objections to the viability of Strong emergence.
- In Chs. 5–8, I consider whether complex systems, ordinary objects, consciousness, and free will, respectively, are actually either Weakly or Strongly metaphysically emergent. For each of these phenomena, I argue that some cases of the phenomenon are plausibly Weakly emergent. For most of these phenomena, I argue that existing arguments for the phenomenon's being Strongly emergent don't go through (though in some cases this remains a live empirical possibility). One exception: I argue that there is presently good reason to think that there is libertarian free will of a Strongly emergent variety.
- In Ch. 9, I finish up and point towards work remaining to be done.

dictability, but only by simulation (see Newman 1996, M. Bedau 1997), lack of conceptual or representational entailment (see Chalmers 1996, Van Gulick 2001), and the presence of theoretical/mathematical singularities (see Batterman 2002).

[4] As I observe, although the assumption that the base-level entities and features are physical or physically acceptable is typically operative in what follows, the schemas generalize to characterize emergence of two different varieties, whatever the precise ontological status of the base-level goings-on.

Besides motivating the book project and setting out the chapter structure, in Ch. 1 I expand on certain suppositions and operative notions informing my investigations (§1.4). In brief:

- *Certain core suppositions*. Notwithstanding their diversity, accounts of metaphysical emergence typically agree on the following theses, which are preserved on my account(s):
  - Metaphysical emergence couples cotemporal material dependence (hence, in particular, does not involve any new substance of the sort posited, e.g., by Cartesian dualists) and some degree of autonomy, where the autonomy at issue is causal as well as ontological.[5]
  - The metaphysical emergence of entities can be investigated by attention to the metaphysical emergence of features of the entities, with the supposition being that if some entity is metaphysically emergent, this is due to its having some characteristic metaphysically emergent feature (e.g., *being conscious*, *being in the basin of a strange attractor*) which can be the target of investigation.
  - Metaphysically emergent features 'minimally nomologically supervene' on base features, in that in every world (actual or hypothetical) with the same or relevantly similar laws of nature, the occurrence of an emergent feature *S* requires the occurrence of some or other base feature *P*, and in every such world, the occurrence of any such *P* will be accompanied by the occurrence of such an *S*.
- *The physical*. Discussions of metaphysical emergence as actually instantiated typically suppose that dependence base goings-on are ultimately physical. But what is it for some goings-on to be physical? The account operative here is that I advance in Wilson 2006, according to which the physical goings-on are those which are treated approximately accurately by present or future (in the limit of inquiry, ideal) physics, with the proviso that the physical goings-on are not fundamentally mental—that is, do not individually either have or bestow mentality. Not much turns on the specific details of the account of the physical, however; the main take-home point is that there is at least one physics-based account of the physical up to the task of characterizing the views at issue.
- *The individuation of levels*. It is common to think of metaphysical emergence in the target cases as going hand-in-hand with the suggestion that emergent entities and features are 'higher-level' with respect to the 'lower-level' goings-on upon which they depend.[6] But which entities and features should be taken to exist at a given level? An important constraint here is that levels (or the one level, if anti-realism or reductionism turns out to be correct) be individuated so as to include any combinations or configurations of entities and features to which the anti-realist or reductionist may reasonably

[5] Even with respect to these components there is some dispute; such variations, however, are either subsumable under the core understandings (as I argue is the case for diachronic accounts of metaphysical emergence; see also Wilson forthcoming*b*) or else are not to the point of accommodating the target phenomena (hence I put aside epiphenomenalist approaches to metaphysical emergence).

[6] Note that 'emergent' and 'higher-level' are not synonymous, however, since non-emergentist views (e.g., Cartesian dualism) also aim to accommodate leveled structure.

appeal. For example, if the basic physical entities are atoms and the basic physical relations include spatial relations and pairwise atomic bonding relations, then we should allow as existing, at the atomic level, not just small numbers of atoms standing in atomic relations, but also large numbers of atoms standing in highly complex atomic (including spatial) relations, constituting pluralities or aggregates of the sort that might, if reductionism is correct, be identical with a rock, a plant, or a person, at least at any given time.

Given this constraint, I offer two different approaches to answering the question of which combinations of entities and associated features should be taken to exist at a given level *L* of broadly scientific reality, beyond the entities and features typically taken, by lights of the associated science *S*, to be characteristic of *L*:

a. *The lightweight combination approach*. Here the individuation of levels proceeds by allowing that various ontologically 'lightweight' (including lower-level relational, mereological, and Boolean) combinations of the characteristic entities and features treated by a given science *S* and placed at a level *L* are also appropriately placed at *L*. For example, the goings-on at the atomic level would include not just atoms and pairwise atomic relations, but any configurations of atoms standing in atomic relations, any boolean combinations of such configurations, and so on.

b. *The 'law-consequence' approach*. Here the individuation of levels proceeds by allowing that any consequences of laws operating at a given level *L*, upon which those laws can operate (take as input), are also appropriately placed at *L*. For example, the goings-on at the atomic level would include any atomic configurations which the atomic laws are capable of taking as input (operating on).[7]

- *The fundamental*. Both physicalists and their Strong emergentist rivals suppose that there are fundamental physical goings-on; where they disagree is over whether there are any fundamental non-physical goings-on. But what is it for some goings-on to be fundamental (at a world, here and throughout)? There are three main approaches (see Tahko 2018 for discussion). On independence-based accounts, what makes it the case that some goings-on are fundamental is that those goings-on are (individually) metaphysically independent. On dependence-based accounts, this is a matter of the goings-on being part of a complete minimal dependence basis for everything that exists. And on primitivist accounts, this is a primitive matter, not metaphysically analyzable in any other terms. (Nota bene that it is not the fun-

[7] Note that on a law-consequence approach, only those consequences of laws at a given level *L* preserving the information required for the *L*-level laws to operate are placed at *L*. As such, a law-consequence approach does not automatically rule out Weak emergence, notwithstanding that Weak emergentists typically maintain that Weak emergents are in some sense metaphysical consequences of physical laws and conditions. For (as an empirical matter—so Weak emergentists argue) the metaphysical consequences associated with Weak emergents typically abstract away from certain lower-level details (e.g., quantum spin) such that were these input into the physical laws, the laws would not have all the information needed for them to operate.

damenta themselves, but what makes it the case that some goings-on are fundamental, that is on these accounts taken to be primitive). My own preference is for a primitivist account, as advanced in my 2014 and developed and defended in my forthcoming*a* and under contract. For the most part, which account of fundamentality is at issue won't matter for what follows, with one exception—namely, an independence-based conception on which individual fundamenta are metaphysically independent (see, e.g., Schaffer 2009, 373; Bennett 2017, 138) rules out fundamenta that are partly but not completely metaphysically dependent on other fundamenta, and so rules out a common understanding of Strongly emergent phenomena. That said, a collectivist variation on an independence-based account, on which the fundamental goings-on collectively do not depend on any other goings on, can accommodate Strong emergence, and so (versions of) all three approaches are suitable for present purposes.

- *Causes and powers.* The discussions to come often advert to causal relations and associated powers to produce effects. More specifically, the schemas for metaphysical emergence that I offer encode certain relations between powers of emergent and dependence base features. There are vast literatures on causation and powers, and on how these notions enter, metaphysically and modally, into the characterizations of entities and features. Fortunately, it is possible to remain almost entirely neutral as regards these more specific details.

  To start, the operative notion of 'power' in what follows is metaphysically highly neutral, following the presuppositions operative in my 2015*b*:

  > [T]alk of powers is simply shorthand for talk of what causal contributions possession of a given feature makes (or can make, relative to the same laws of nature) to an entity's bringing about an effect, when in certain circumstances. That features are associated with actual or potential causal contributions ('powers') reflects the uncontroversial fact that what entities do (can do, relative to the same laws of nature) depends on how they are (what features they have). So, for example, a magnet attracts nearby pins in virtue of being magnetic, not massy; a magnet falls to the ground when dropped in virtue of being massy, not magnetic. Moreover, a feature may contribute to diverse effects, given diverse circumstances of its occurrence (which circumstances may be internal or external to the entity possessing the feature). Anyone accepting that what effects a particular causes (can cause, relative to the same laws of nature) is in part a function of what features it has—effectively, all participants to the present debate—is in position to accept powers, in this shorthand, metaphysically neutral and nomologically motivated sense (354).

  The operative notion of causation is also highly metaphysically neutral. By way of proof of concept, I argue that even a contingentist categoricalist Humean—someone who maintains that causation is a matter of regularities, features have their powers contingently, and all features are ultimately categorical—can accept powers and the associated notion of causation in the neutral sense(s) here. For such a Humean, to say that an (ultimately categorical) feature has a certain power would be to say that, were a token of the feature to occur in certain circumstances, a certain (contingent) regularity would be instanced.

More generally, no controversial theses pertaining to the nature of powers, causation, properties, or laws are presupposed in the discussions or the schemas to follow. That said, I do suppose that we can make sense of physical causation. Some (e.g., Russell, 1912, and Field, 2003) claim that this is problematic; but first, the Russell/Field position is an outside view, as is clear from the usual formulations of physicalism as committed to Physical Causal Closure, according to which any physical effect has a sufficient purely physical cause; second, in any case, I argue that the Russell/Field line(s) of thought can be resisted.

- *Methodology*. Following most contemporary metaphysicians, I implement a broadly abductive methodology (i.e., 'inference to the best explanation', per Harman 1965 and Douven 2021), whereby candidate metaphysical accounts of a given phenomenon are assessed by attention to how well they do, overall, at satisfying various theoretical desiderata. To be sure, there is variation in exactly which theoretical desiderata are operative as well as in how these desiderata, which may push in different directions, should be weighted. As I discuss in my 2011, 2016*c*, and 2016*b*, this variation is unsurprising, given the wide purview of metaphysical investigations and our present distance from the end of inquiry. Even in the absence of complete consensus regarding methodological standards, progress can be made, so long as one is suitably explicit about which theoretical desiderata are primarily guiding one's investigations. Two methodological desiderata which I take to be especially important in my theorizing are as follows:
    1. *Criterion of Appropriate Accommodation*: An adequate account of metaphysical emergence should make natural (straightforward, default) and realistic sense of the appearances of metaphysical emergence, in the absence of reasons to think that this cannot be done. Hence while I take it to be part of my burden to show that various purported problems with metaphysical emergence can be addressed, I do not take it to be part of my burden to show that no deflationary (anti-realist or reductionist) account of the appearances of metaphysical emergence is viable. My ultimate goal is not to knock the anti-realist or reductionist off their horse, but to show the metaphysical emergentist who aims to accommodate the appearances at realistic face value how to stay on their own horse. I hope that those with different methodological sensibilities will nonetheless find the ensuing discussion useful, at least as revealing the extent to which the heavy weighting of parsimony considerations, as opposed to any specific problem with the notion of metaphysical emergence itself, may be playing a role in deflationary accounts of such emergence.
    2. *Criterion of Illuminating Accommodation*: An adequate account of metaphysical emergence should provide an illuminating basis for accommodating the appearances of metaphysical emergence in natural (straightforward, default) fashion. Hence it isn't enough to simply stipulate, or take it to be brute or primitive, that some goings-on are both cotemporally materially dependent and suitably autonomous; what is desired is one or more intelligible, explanatory account(s) of how there can be metaphysical emergence in this sense.

Chapter 2: "Two Schemas for Metaphysical Emergence"

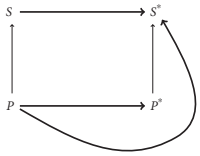
In Chapter 2, I motivate my two schemas for metaphysical emergence by attention to what is seen by many as the most pressing challenge to taking the appearances of metaphysical emergence as genuine—namely, the problem of higher-level causation, made salient by Kim in his 1989, 1993a, 1998, and elsewhere. I argue, following discussions in Wilson 1999, 2001, 2011*b*, and elsewhere, that there are two and only two strategies of response to this problem that make sense of seemingly higher-level entities and features' being metaphysically emergent as above. One strategy provides a schematic basis for 'Weak' (physically acceptable) emergence; the other provides a schematic basis for 'Strong' (physically unacceptable) emergence.[8] For each of these strategies and associated schemas, I show that a representative range of seemingly diverse accounts of metaphysical emergence are plausibly seen as satisfying the conditions in one or the other schema, and thus are more unified than they appear.

I start by presenting Kim's problem of higher-level causation (§2.1). The general concern is that any purported effects of higher-level features are already produced by the lower-level features upon which they minimally nomologically supervene, such that the metaphysical emergentist is committed to such effects' being problematically causally overdetermined—that is, problematically caused twice over. More specifically, the problem is usefully seen as involving the following six premises:

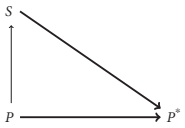
1. *Dependence*. Special science features cotemporally materially depend on lower-level physical features ('base features').
2. *Reality*. Both special science features and their base features are real.
3. *Efficacy*. Special science features are causally efficacious.
4. *Distinctness*. Special science features are distinct from their base features.
5. *Physical Causal Closure*. Every lower-level physical effect has a purely lower-level physical cause.
6. *Non-overdetermination*. With the exception of cases of the double-rock-throw variety, effects are not causally overdetermined by distinct individually sufficient cotemporal causes.

There are two cases to consider, reflecting two sorts of effect. In Kim's presentation, *S* is a mental state (e.g., *being thirsty*); *P* is a base state upon which *S* depends; and *S* is taken to cause either another mental state *S** (e.g., a desire to quench one's thirst) or a base state *P** (e.g., a physical reaching for a glass of water). But the challenge more generally concerns how any real, distinct, dependent higher-level feature might be unproblematically efficacious. The two cases are as follows (bold lines = causation, thin lines = cotemporal material dependence):

[8] Again, the schemas more generally operate to characterize emergence of two different varieties, whatever the precise ontological status of the base-level goings-on.



Case 1 of the problem of higher-level causation: *S* causes *S**



Case 2 of the problem of higher-level causation: *S* causes *P**

Kim rejects *Distinctness*, favouring reductive physicalism. But more generally (see Wilson 2015), rejection of each premise is associated with certain prominent views. To start:

1. *Substance dualism*. Deny *Dependence*: avoid overdetermination by denying that *S* and *S** cotemporally materially depend on base features *P* and *P**, respectively.
2. *Eliminativism*. Deny *Reality*: avoid overdetermination by denying that *S* and *S** are real.
3. *Epiphenomenalism*. Deny *Efficacy*: avoid overdetermination by denying that *S* is efficacious.
4. *Reductive physicalism*. Deny *Distinctness*: avoid overdetermination by identifying *S* with *P*.

These strategies avoid overdetermination, but don't make sense of higher-level features as metaphysically emergent—that is, as real, dependent, distinct, and distinctively efficacious.

There are, however, two strategies of response to Kim which do accommodate metaphysical emergence:

5. *Strong emergentism*. Deny *Physical Causal Closure*: avoid overdetermination by denying that every lower-level physical effect has a purely lower-level physical cause. This is the strategy encoded in 'British Emergentist' accounts.
6. *Weak emergentism*. Deny *Non-overdetermination*: allow that effects caused by *S* are also caused by *P*, but maintain that the overdetermination here is of an unproblematic *non*-double-rock-throw variety. This is the strategy encoded in non-reductive physicalist accounts (e.g., functional realization, determinable-determinate, and constitutive mechanism accounts).

As I argue in the next two sections, these two strategies and associated positions are perspicuously seen as motivated by two conditions on the powers of a given special-science feature, where satisfaction of one or other condition provides a prima facie plausible and principled (i.e., appropriate and illuminating) basis for taking the feature to be emergent, in ways that standard proponents of the strategy/position would endorse. In each of these sections, treating Strong emer-

gence and Weak emergence, respectively, I start by motivating the associated condition on powers by attention to standard versions of the position; I then show how satisfaction of the condition dovetails with the associated strategy for responding to the problem of higher-level causation; I then provide prima facie reasons for thinking that satisfaction of the condition provides an appropriate and illuminating basis for taking special-science features to be both cotemporally materially dependent and ontologically and causally autonomous; finally, I use the condition to formulate the associated schema for metaphysical emergence.

The Schema for Strong Emergence

I start with the Strong emergentist strategy, as implemented most saliently by British emergentists (§2.2). The conception of higher-level efficacy at issue in Strong emergentism is, as above, one which denies *Physical Causal Closure*, and is correspondingly incompatible with physicalism. And while different accounts of Strong emergentism emphasize different aspects of this distinctive efficacy as located in fundamentally novel features, laws, effects, forces, or interactions, core and common to these accounts is that Strongly emergent features have fundamentally novel powers—powers to produce effects entailing the violation, in particular, of *Physical Causal Closure*, as per the following condition:

> *New Power Condition*: Token feature *S* has, on a given occasion, at least one token power not identical with any token power of the token feature *P* upon which *S* cotemporally materially depends, on that occasion.

This is true, to start, on British emergentism, as endorsed most systematically by Mill (1843/1973), Alexander (1920), Lewes (1875), and Broad (1925). Hence in his classic survey, McLaughlin (1992) describes British emergentism as

> [T]he doctrine that there are fundamental powers to influence motion associated with types of structures of particles that compose certain chemical, biological, and psychological kinds" (52), where the powers at issue are typically taken to be "powers to generate fundamental forces not generated by any pairs of elementary particles. (71)

Contemporary accounts of Strong emergence also typically agree in taking emergent features to have or bestow fundamentally novel powers, not had (or had only in derivative fashion) by base features or associated microconfigurations. For example, O'Connor and Wong (2005) characterize emergent features as "fundamentally new", not just in being (perhaps epiphenomenally) different, but more specifically in having fundamentally novel causal capacities:

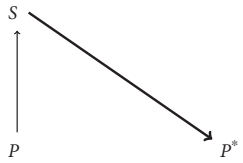
> [A]s a fundamentally new kind of feature, [an emergent feature] will confer causal capacities on the object that go beyond the summation of capacities directly conferred by the objects microstructure. (665)

(See also, e.g., Silberstein and McGeever 1999, Wilson 1999, and Van Gulick 2001.)

Given that higher-level feature *S* has a (fundamentally novel) power to cause a given effect—a power that its dependence base feature *P* does not

have—the Strong emergentist's responses to Kim's cases can be represented as follows:

The Strong emergentist's response to case 1



The Strong emergentist's response to case 2

Prima facie, satisfaction of the New Power Condition by a special-science feature *S* which cotemporally materially depends on a base feature *P* provides an appropriate and illuminating basis for avoiding overdetermination while guaranteeing that *S* is both ontologically and causally autonomous with respect to *P*. We have thus arrived at our first schema for metaphysical emergence:

> *Strong Emergence*: What it is for token feature *S* to be Strongly metaphysically emergent from token feature *P* on a given occasion is for it to be the case, on that occasion, (i) that *S* cotemporally materially depends on *P*, and (ii) that *S* has at least one token power not identical with any token power of *P*.

Here the locution 'what it is for' is intended to flag that Strong Emergence provides a schematic metaphysical basis for a given case of such emergence, encoding what is core and crucial to that notion. Some clarifications:

- The notion of 'power' operative in the schema is metaphysically highly neutral.
- The base feature *P* in the schema is a feature of a micro-configuration (not of an individual component of the configuration), and the conditions should be understood accordingly.
- The first condition encodes substance monism and minimal nomological supervenience.
- The second condition ensures ontological and causal autonomy (distinctness and distinctive efficacy). For Strong emergence, distinctive efficacy involves the higher-level feature's having *a new power*—a power not had, or not had in same way, by the base feature:
  - Note that the novel token power is fundamentally novel, since non-fundamentally novel powers (powers had just in virtue of aggregation) are had by base feature *P*.
  - In having a novel token power, *S* can cause an effect that *P* can't cause, or that *P* can't cause in the same (non-derivative) way as *S*;

hence *S* is causally autonomous—that is, distinctively efficacious—with respect to *P*.

– That a Strong emergent has a token power not had by its base feature *P* entails that *S* is distinct from *P*, by Leibniz's Law.

- The schema is relativized to occasions (times or temporal intervals), but it would be reasonable to suppose that it suffices for the Strong emergence of *S*, simpliciter, that the condition is ever satisfied, and to suppose that it suffices for the Strong emergence of the feature type (of which *S* is a token), simpliciter, that any token feature *S* on any occasion satisfies (or would satisfy) the condition.

The Schema for Weak Emergence

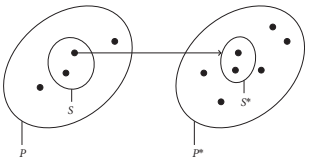
I focus next on the Weak emergentist strategy, as implemented most saliently by non-reductive physicalists (§2.3). Like Strong emergentists, non-reductive physicalists maintain that (some) higher-level features are real, cotemporally materially dependent, distinct, and distinctively efficacious with respect to their base features. But as physicalists, their response to the problem of higher-level causation cannot entail the rejection of *Physical Causal Closure*, which is core to the physicalist view that the physical goings-on are an existential and causal basis for all other broadly scientific phenomena. Rather, non-reductive physicalists reject *Non-overdetermination*, maintaining that distinct special science and base features can each be sufficient causes of a single effect, in virtue of standing in a relation that, while not identity, is intimate enough both to avoid overdetermination of the problematic (since implausible, for the cases at issue) double-rock-throw variety and to retain compatibility with *Physical Causal Closure*, hence with physicalism.

Non-reductive physicalists posit a variety of relations as showing how it can be that a higher-level feature can be completely metaphysically dependent on, yet distinct and distinctively efficacious with respect to, lower-level dependence base features. These include functional realization (Putnam 1967, Fodor 1974, Papineau 1993, Antony and Levine 1997, Melnyk 2003, Witmer 2003, Polger 2007, Yates 2012), the determinable-determinate relation (MacDonald and MacDonald 1986, Yablo 1992, Wilson 1999 and 2009), constitutional mechanism (Cummins 1975, Craver 2001, Haug 2010), mereological realization (Shoemaker 2000/2001, Clapp 2001, Rueger and McGivern 2010), and many others. Though there are interesting differences between these accounts of non-reductive realization, I argue that they have in common that each is plausibly such as to satisfy the following condition on token powers of realized and realizing features:

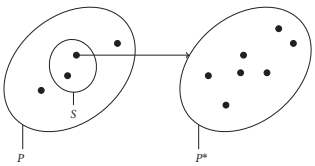
Proper Subset of Powers Condition: Token feature *S* has, on a given occasion, a non-empty proper subset of the token powers of the token feature *P* on which *S* cotemporally materially depends, on that occasion.[9]

Representing the features at issue as having overlapping sets of powers, with each power represented as a dot, the non-reductive physicalist's responses to Kim's cases are as follows:

[9] The requirement that the proper subset of powers be non-empty reflects the rejection of epiphenomenal features as metaphysically emergent, in the relevant sense.



The Weak emergentist's response to case 1



The Weak emergentist's response to case 2

Prima facie, satisfaction of the Proper Subset of Powers Condition by a special-science feature *S* which cotemporally materially depends on a base feature *P* provides an appropriate and illuminating basis for avoiding overdetermination while guaranteeing that *S* is both ontologically and causally autonomous with respect to *P*. We have thus arrived at our second schema for metaphysical emergence:

> *Weak Emergence*: What it is for token feature *S* to be Weakly metaphysically emergent from token feature *P* on a given occasion is for it to be the case, on that occasion, (i) that *S* cotemporally materially depends on *P*, and (ii) that *S* has a non-empty proper subset of the token powers had by *P*.

Here again, the locution 'what it is for' is intended to flag that Weak Emergence provides a schematic metaphysical basis for a given case of such emergence, encoding what is core and crucial to that notion. Some clarifications:

- The notion of 'power' operative in the schema is metaphysically highly neutral, as is the supposition that one can make sense of the identity (non-identity) of powers (see my reply to Bennett for further discussion).
- The base feature *P* in the schema is a feature of a micro-configuration (not of an individual component of the configuration), and the conditions should be understood accordingly.
- The first condition encodes substance monism and minimal nomological supervenience.
- The second condition ensures ontological and causal autonomy (distinctness and distinctive efficacy. For Weak emergence, distinctive efficacy involves the higher-level feature's having *strictly fewer* powers than are had by the base feature, and hence having a distinctive power profile:
  - Here the response to Kim proceeds by maintaining—contra what Kim assumes—that distinctive efficacy of a higher-level feature does not require that it have a new power.
  - It suffices for distinctive efficacy that the feature have a distinctive power profile, tracking difference-making considerations (if my thirst had been differently physically realized, I would still have reached for

the Fresca), or comparatively abstract levels of causal or nomological grain.

- That a Weak emergent has a distinctive power profile entails that it is distinct from its base feature, by Leibniz's Law.
- Again, the schema is relativized to occasions (times or temporal intervals), but it is reasonable to suppose that (given that S's type is not Strongly emergent) it suffices for the Weak emergence of S, simpliciter, that the condition is ever satisfied, and to suppose that it suffices for the Weak emergence of the feature type (of which S is a token), simpliciter, that any token feature S on any occasion satisfies (or would satisfy) the condition.

I close the chapter by observing that attention to the problem of higher-level causation makes clear the limited ways in which a cotemporally materially dependent higher-level feature can be causally, hence ontologically, autonomous with respect to its base feature, as the operative conception of metaphysical emergence requires (§2.4). First, the feature may have *more* powers than its base feature, as in Strong emergence;[10] second, the feature may have fewer powers than its base feature, as in Weak emergence. In terms of effects: the higher-level feature may be distinctively efficacious in potentially contributing to causing at least one different effect than its base feature (Strong emergence), or it may be distinctively efficacious in potentially contributing to fewer effects than its base feature (Weak emergence). Since complete coincidence of token powers doesn't make room for causal autonomy (distinctive efficacy), these routes to metaphysical emergence exhaust the available options.

I conclude that satisfaction of the conditions in either schema is, as I put it, 'core and crucial' to metaphysical emergence of the sort relevant to realistically vindicating the seeming appearances of emergence as pertaining to special-scientific and artifactual entities and features. Modulo the supposition that the schemas are sensibly filled in, the results of this chapter can be seen as providing prima facie reason to think that the conditions in the schemas are both necessary and sufficient for (appropriate and illuminating accommodation of) metaphysical emergence of both physically acceptable and physically unacceptable varieties—a bold claim, but one that, as I argue in ensuing chapters, is surprisingly robust.

Chapter 3: "The Viability of Weak Emergence"

In Chapter 3, I consider and respond to a representative range of objections to the viability of Weak emergence, understood as per the associated schema:

> *Weak Emergence:* What it is for token feature S to be Weakly metaphysically emergent from token feature P on a given occasion is for it to be the case, on that occasion, (i) that S cotemporally materially depends on P, and (ii) that S has a non-empty proper subset of the token powers had by P.

These objections fall into four main categories, according to which satisfaction of the conditions in Weak Emergence is …

[10] By 'more' I just mean that a Strong emergent must have at least one power not had by the base feature; pace Ney (2022), I do not suppose (and nor does satisfaction of the conditions in the schema require) that a Strong emergent have all the powers of the base feature, and then some.

- compatible with anti-realism about higher-level features (§3.1);
- compatible with reductionism about higher-level features (§3.2);
- compatible with the emergent feature's being physically unacceptable (§3.3); or
- not necessary for metaphysical emergence of a physically acceptable variety (§3.4).

The primary focus of many of the objections is on condition (ii) in the schema—i.e., the Proper Subset of Powers Condition. These diverse challenges can, I argue, be answered. Each of these objections admits of at least one response that could be endorsed by any proponent of Weak emergence, whatever their preferred implementation of the schema. Upon occasion, however, I offer certain attractive responses appealing to either a determinable-based account of Weak emergence (per my 1999 and 2009, developing the proposals in MacDonald and MacDonald 1986 and Yablo 1992), or an account of Weak emergence as involving an elimination in degrees of freedom (per my 2010, developing the proposal in Batterman 1998 and elsewhere).

Here, by way of partial illustration, I sketch certain representative lines of response to each of the four categories of concern.

According to the first concern (see, e.g., Heil 2003, Ney 2010, and Morris 2018), "nothing has been said to rule out" (as Ney puts it) an abstractionist or pragmatist line on seeming satisfaction of the Proper Subset of Powers Condition. I grant that this is the case, but deny that the viability of Weak emergence hinges on accomplishing such a 'ruling out.' Given the many prima facie reasons for thinking that there is metaphysical emergence, the burden is on the anti-realist to provide reasons for not taking the appearances at face value; but so far anti-realists have not provided any such good reason—in particular, as telling against a Weak emergentist treatment of the appearances. For example, Heil suggests that predicates such as 'red' should be understood not as referring to higher-level features, but rather as tracking inexact similarities between lower-level features, especially in light of Kim-style overdetermination concerns; but even granting that the predicates at issue are tracking inexact similarities among lower-level features, this would not show that the higher-level features did not exist, unless it was antecedently clear that the inexact similarities at issue were not themselves higher-level, which it isn't; and as above, the Weak emergentist has a response to Kim's overdetermination concerns, which makes clear how Weak emergents can be causally efficacious in spite of not having any new powers, in virtue of having a distinctive power profile, tracking difference-making considerations and comparatively abstract levels of causal grain.

According to the second concern, even granting that feature *S*'s satisfying the conditions in Weak emergence physical feature *P* ensures that *S* is real and distinct from *P*, this much is compatible with *S*'s being ontologically reducible to—that is, identical with—some *other* lower-level physically acceptable feature P' (see Yates 2012, 6, for discussion of the general concern). There are diverse reductive strategies here, according to which *S* is reducible to …

- a conjunct of a lower-level conjunction (§3.2.1);
- a disjunction of lower-level disjuncts (§3.2.2); or
- a metaphysical consequence of lower-level laws (§3.2.3).

To each strategy I offer one or more responses that any Weak emergentist might accept. In the case of the first strategy (see Shoemaker 2000/2001 for discussion), one might stipulatively rule out conjunctive realization (as Shoemaker does), or implement Baysan's suggestion that, on the supposition that conjunct features are more fundamental than associated conjunctive features, a conjunct feature *S* would not be appropriately taken to satisfy the relevant condition on dependence in the schema for Weak emergence. I additionally note that an appeal to a determinable-based implementation of Weak emergence will suffice to non-stipulatively rule out conjunctive realization, since it is definitive of the determinable/determinate relation that it is not properly metaphysically characterized in terms of anything like the conjunct/conjunction (or relatedly, genus/species) relations (see Wilson 2022/2017 for discussion). In the case of the second 'disjunctive' strategy (see, e.g., Fodor 1987, Jaworski 2002, and Dosanjh 2014 and 2019), I argue that on the usual understanding according to which what it is for a disjunctive type to be tokened on a given occasion is for one of the disjunct types to be tokened on that occasion, the disjunctive strategy is incompatible with satisfaction of the Proper Subset of Powers Condition. And in the case of the third strategy (see, e.g., Nagel 1961, Klee 1984, Kim 2010, and Morris 2018), I observe (see note 7 of this précis) that a proper understanding of how laws enter into the individuation of levels enables the Weak emergentist to maintain that, notwithstanding that special scientific goings-on are, on their view, metaphysical consequences of lower-level physical goings-on, it does not follow that the former are identical with any of the latter, since the former do not contain all the information needed for the lower-level physical laws to operate. I additionally note that a DOF-based implementation of Weak emergence develops this idea, in that on this implementation special-science goings-on may be metaphysical (and even deductive, so to speak) consequences of lower-level physical goings-on, yet be distinct from any lower-level physical goings-on, in failing to have all the DOF that are needed for the lower-level physical laws to operate (as first discussed in Wilson 2010).

According to the third line of concern, that a feature *S* satisfies the conditions in Weak emergence vis-à-vis a given physical feature *P* is compatible with *S*'s being physically unacceptable. Again, there are several variations of the theme of the concern, according to which satisfaction of the Proper Subset Condition on Powers, in particular, is compatible with *S*'s being 'over and above' *P* in virtue of …

- *S*'s having a non-causal quiddity (§3.3.1);
- *S*'s having a phenomenal aspect (§3.3.2);
- S's failing to be entailed by *P* (§3.3.3);
- *S*'s having a fundamentally mental power (§3.3.4); or
- *S*'s being associated with physically unacceptable constraints (§3.3.5).

In re non-causal quiddities (per Melnyk 2006, Morris 2018), I argue that the Weak emergentist can reasonably maintain that whether *S* and/or *P* have quiddities, shared or not, is irrelevant to whether *S* is physically acceptable, since the occurrence of scientific features, and any truths about such features, does not depend on or otherwise track whether such features have quiddities, much less track how the noncausal quiddities of seemingly distinct features are related; and similarly for artifactual features satisfying the conditions in Weak Emer-

gence. In re phenomenal aspects (per, e.g., Walter 2010), I argue that the common supposition that phenomenal aspects (of mental features, in particular) cannot be characterized in terms of causal roles or associated powers is incorrect; rather, as per what I call the 'Phenomenal Incorporation Thesis,' phenomenal aspects of mental features are fully incorporated into the powers of these features (compatible with powers' being contingently associated with features, relative to a given set of laws), reflecting that differences in phenomenality give rise to causal differences. In re a supposed failure of *S* to be entailed or necessitated by *P* (per Melnyk 2006, McLaughlin 2007), I observe (among other responses) that the cases usually offered as showing that *S* would be 'over and above' *P* in not even being nomologically entailed or necessitated by *P* fail to take the cotemporal material dependence condition in Weak emergence into account. In re fundamentally mental powers (per Baltimore 2013), I observe that while the Proper Subset Condition on Powers itself does not rule out *P*, hence *S*, from having fundamentally mental powers, the operative 'no fundamental mentality' account of the physical (per my 2006) does so. Finally, in re physically unacceptable constraints (per Melnyk 2006), I grant that when the Proper Subset Condition is satisfied as a result of constraints being imposed on lower-level goings-on, the constraints themselves need to be physically acceptable, and that it might be worth adding this requirement to the schema for Weak emergence (as I explicitly do in my DOF-based implementation of Weak emergence).

According to the fourth line of concern, satisfaction of the conditions in Weak emergence is not necessary for physically acceptable emergence; rather, one or other account in terms of token identity (per Davidson 1970, Macdonald and Macdonald 1995, Ehring 2003, and Robb 1997) (§3.4.1), constitutive mechanism (per Gillett 2002*a*, 2002*b*, 2016) (§3.4.2), constitution (per Pereboom 2002) (§3.4.3), or primitive Grounding (per Schaffer 2009, Rosen 2010, and Dasgupta 2014) (§3.4.4) will do the job. Considerations of space prevent my discussing these alternatives in any detail here; I can say, however, that a common theme is that the views at issue either fail to establish the ontological and causal autonomy of higher-level features, and so are not really accounts of physically acceptable emergence; or else are plausibly seen as imposing the Proper Subset of Powers Condition, and so are not really competitors to my view.

Chapter 4: "The Viability of Strong Emergence"

In Chapter 4, I consider and respond to a representative range of objections to Strong emergence, understood as per the associated schema:

> *Strong Emergence:* What it is for token feature S to be Strongly metaphysically emergent from token feature P on a given occasion is for it to be the case, on that occasion, (i) that S cotemporally materially depends on P, and (ii) that S has at least one token power not identical with any token power of P.

These objections fall into four main categories, according to which satisfaction of the conditions in Strong Emergence is …

- incompatible with scientific theory or practice (§4.1);
- impossible, since any purportedly novel powers of Strongly emergent features are inherited by (or 'collapse' into) base features (§4.2);
- compatible with physical acceptability (§4.3); or

- not necessary for emergence of a physically unacceptable variety (§4.4).

Here again, I argue that these diverse challenges can be answered. And here again, each objection admits of at least one response that any proponent of Strong emergence could endorse, whatever their preferred implementation of the schema. Upon occasion, however, responses draw on features of my preferred 'fundamental interaction-relative' account of Strong emergence (as per my 2002), according to which a Strongly emergent entity (feature) has at least one power that is grounded, at least in part, in a novel (nonphysical) fundamental interaction.

Here, by way of partial illustration, I sketch certain representative lines of response to each of the four categories of concern.

According to the first commonly voiced concern, Strong emergence is naturalistically or scientifically unacceptable. In response, I start by observing, following McLaughlin 1992, that Strong emergence would not be incompatible with laws such as $F = ma$ or Schrödinger's equation, but would rather just involve adding another force or energy to the mix of those input into these laws of nature. I moreover argue, following Wilson 2002, that reflecting that scientific practice suggests that powers are plausibly grounded, one way or another, in fundamental forces or interactions (as when the power of a magnet to attract a pin is grounded in the electromagnetic interaction), naturalistic good sense can be made of the Strong emergentist posit of fundamentally novel powers, as reflecting novel fundamental interactions that come into play only at certain levels of compositional complexity, such that Strong emergentism "is committed to there being at least one other fundamental force beyond those fundamental forces currently posited" (74). Indeed, the case of the weak nuclear interaction, posited in response to apparent conservation law violations in beta decay, supports the naturalistic/scientific respectability of Strong emergence: since a nucleus is a complex entity, evidently scientists have no problem with positing fundamental configurational interactions and associated powers. Similar experiments could provide an empirical basis for Strong emergence, in principle.

Finally, I observe that claims that there is "not a scintilla of evidence" in favor of there being Strongly emergent features (McLaughlin 1992; see also Ladyman and Ross 2007) are overstated, especially in light of the result forthcoming in Ch. 8 (see also my response to McLaughlin, this volume).

According to the second concern, Strong emergence is impossible, due to the base feature's inheriting any purportedly novel power, as per what Taylor (2015) evocatively calls the 'collapse' objection (see Cleve 1990, Kim 1999, O'Connor 1994, Wilson 2002, Francescotti 2007, Howell 2009, Taylor 2015, and Carruth 2018). Drawing on Baysan and Wilson 2017, I offer four strategies for avoiding collapse. Three might be implemented by any account of Strong emergence; these involve (i) distinguishing between direct and indirect having of powers, (ii) distinguishing between lightweight and heavyweight dispositions, and (iii) taking Strongly emergent features to be 'new object entailing,' in ways that block lower-level inheritance of powers. The fourth strategy draws on my fundamental interaction-relative account of Strong emergence. On this account, to start, powers are grounded (I make some specific suggestions as to how) in fundamental interactions: as above, magnets have the power to attract pins in virtue of the electromagnetic, not the gravitational, interaction; and so on. One can understand the New Power Condition accordingly. Relative to the set of

purely physical fundamental interactions, a cotemporally materially dependent feature S can have a fundamentally novel power p, as per the schema for Strong emergence; relative to the set of any and all fundamental interactions, p will be inherited by the lower-level physical features P upon which S cotemporally materially depends.

According to the third concern (due to Yates 2016), satisfaction by a feature S of the conditions in Strong emergence is compatible with S's being physically realized, hence physically acceptable. By way of illustrative motivation Yates argues that the molecular geometry G of a water molecule is a mathematically specified, physically realized feature which bestows certain powers upon its bearer—in particular, those, including hydrogen bonding in water, associated with the molecule's dipole moment—not had/bestowed by G's realizers. Here I argue that Yate's reasons for thinking that the powers had by G are not had by the base feature F that 'qualitatively' realizes G on a given occasion do not go through. In particular, he supposes that if such power inheritance were in place, references to G could be eliminated in broadly deductive explanations of the dipole moment and associated powers, yet such references can't be eliminated; but (I observe) nothing in physicalism or in the physicalist supposition that higher-level features inherit their powers from physical base features requires that elements of higher-level explanations, deductive or otherwise, be 'dischargeable' in terms referring only to lower-level physical goings-on. Moreover, Yates maintains that G can be deduced from lower-level physical goings-on, as an "intermediary step"; but then why think that the need to appeal to G indicates that G has new powers, as opposed to thinking that this need simply reflects that the explanation of the existence and powers of the dipole moment has to proceed in steps, compatible with the physicalist assumption that any powers of deducible features such as G are inherited? More generally, I argue that Yates does not establish that the relation of qualitative realization is (like functional and other forms of realization) also a relation of causal power bestowal.

According to the fourth concern, satisfaction of the conditions in Strong Emergence is not necessary for physically unacceptable emergence. There are four main alternative approaches on offer, in terms of …

- epiphenomenalism (§4.4.1);
- supervenience (§4.4.2);
- primitivism (§4.4.3); or
- epistemic criteria (§4.4.4).

In response, I provide reasons for thinking that each of these alternative approaches to physically unacceptable emergence is unsatisfactory. Again, considerations of space prevent my discussing these alternatives in any detail; here I briefly register some lines of argument.

In re epiphenomenalism (per, e.g., Chalmers 1996): the motivations for making room for an epiphenomenalist conception of emergence rest on there being phenomenal properties, along with the assumption that such properties cannot be characterized in terms of causal roles or associated powers; but as per the 'Phenomenal Incorporation Thesis,' discussed above, this is incorrect. In re supervenience (per, e.g., Chalmers 2006, Witmer 2001): I first canvass reasons for thinking that Strong emergence cannot be characterized as involving nomological but not metaphysical necessity of emergent on base features, since (per sce-

narios highlighted in, e.g., Horgan 1993 and Wilson 2005) Strongly emergent features might supervene with metaphysical necessity on base features. I then offer several responses to Howell's 2009 argument that such scenarios pose no threat to a supervenience-based characterization of such emergence, since metaphysically necessitated features would 'pollute' the dependence base features in such a way that the latter would no longer be properly considered physical, including one according to which (as in the case of a fundamental interaction-based response to the collapse objection) fundamental interactions provide a basis for distinguishing lower-level physical from Strongly emergent goings-on, even when these are deeply dispositionally connected. In re a view on which Strongly emergent goings-on are those which are both fundamental and dependent, and where the notions of fundamentality and dependence are each taken to be primitive (per Barnes 2012): I argue that such a view is too abstract to satisfy the criteria of appropriate and illuminating accommodation; relatedly, it does not provide any clear means of engaging with or addressing either Kim's problem of higher-level causation or the collapse objection, or of ensuring that Strongly emergent goings-on properly contrast with views such as substance dualism. Finally, in re epistemic criteria: I argue that while accounts of Strong emergence as involving one or other epistemic failure have been historically common—per, e.g., appeals to failures of deducibility (Broad 1925), explainability (Horgan 1993), or conceptual entailment (Chalmers 2006), such accounts should be rejected, both because it is clear that the proponents offer the epistemic criteria in service of tracking a metaphysical distinction—in particular, one conforming to the conditions in Strong emergence, and because in any case such epistemic failures are not distinctive of physically acceptable emergence, but can attach to phenomena (e.g., the behaviour of artificial complex systems; see below) for which Strong emergence is clearly not at issue.

Chapter 5: "Complex Systems"

Having established the in-principle viability of both Weak and Strong conceptions of metaphysical emergence, I go on to consider whether certain phenomena are plausibly seen as actually either Weakly or Strongly emergent. I start in Chapter 5 with complex systems, as perhaps the phenomena that have been most often offered as emergent, by scientists as well as philosophers. Complex systems take many forms, both natural (e.g., turbulent water flows, phase transitions, and weather patterns) and artificial (e.g., Conway's 'Game of Life'). And among the distinctive characteristics of complex systems are non-linearity (whereby certain features or behaviours cannot be seen as linear or other broadly additive combinations of features of the system's composing entities), unpredictability (and relatedly, extreme sensitivity to initial conditions), algorithmic incompressibility (whereby the operative equations of motion do not admit of analytic or 'closed' solutions'), 'universality' (whereby certain features are common across diverse micro-structures, especially as associated with asymptotic singularities near critical points), and self-organization (whereby coherent 'system-wide' patterns arise as a result of interactions between parts).

I first consider whether any complex systems might be Strongly emergent (§5.1). I start with a compressed historical discussion of why the British Emergentists (Mill and Broad, among others) took nonlinearity and in-principle failures of predictability to suffice for fundamental novelty (§5.1.1)—a view that,

while reasonable at the time, was undermined by the discovery and creation of complex systems clearly not involving any fundamentally novel powers/interactions/laws. This discussion is useful for appreciating how nonlinearity moved from being a criterion of Strong emergence to being a criterion of Weak emergence (though in ways leaving open, as I argue in §5.1.3, the possibility that some complex systems are Strongly emergent), and for seeing how a recognizable descendant of nonlinearity as a criterion of Strong emergence is present in the aforementioned motivation for new fundamental interactions, reflecting seeming violations of conservation laws. By lights of the latter criterion, I observe, there is presently little support for taking non-mental complex systems to be Strongly emergent (§5.1.4)—though the case is less clear for certain mental phenomena, a topic to which I return in later chapters.

I next consider whether any complex systems might be Weakly emergent (§5.2), focusing on three existing cases for such emergence as involving one or other characteristic of such systems: Bedau's (1997 and 2008) appeal to algorithmic incompressibility (§5.2.1), Mitchell's (2012) appeal to self-organization (§5.2.2), and Batterman's (2000 and 2002) appeal to asymptotic singularities (§5.2.3). I argue that the cases made in these discussions fall short of establishing that complex systems are Weakly emergent, in failing to rule out certain reductionist strategies for accommodating the characteristics at issue. That said, I go on to argue that the prospects for developing these cases in a way that reveals an associated satisfaction of the conditions in Weak Emergence are good (§5.2.4). In particular, after expanding a bit on my (2010) degree-of-freedom (DOF)-based account of Weak emergence, and responding to the concern, due to Morrison (2012) and Lamb (2015), that complex systems involve not fewer but more DOF than base systems (associated with 'order parameters' that emerge near critical points), I argue that complex systems exhibiting universality of the sort Batterman focuses on also have (as he observes) DOF that are eliminated relative to the systems of their composing lower-level entities, and so are Weakly emergent by lights of a DOF-based account. And I go on to offer reasons for thinking that certain other complex systems (Bedau's gliders in Conway's Game of Life; Mitchell's flocks of birds) may also be seen as Weakly emergent by these lights.

Chapter 6: "Ordinary Objects"

In Chapter 6, I turn to the question of whether ordinary objects are either Strongly or Weakly metaphysically emergent. By 'ordinary' objects I have in mind objects which are uncontroversially inanimate (as Thomasson, 2007, puts it) or nonliving (as Merricks, 2003, puts it), and of the sort with which creatures like us are or may be perceptually acquainted. Such objects might be either natural (rocks, feathers, mountains, planets) or artifactual (tables, baseballs, statues). My discussion is broadly neutral on which metaphysical account of objects is correct, so long as a given such account does not rule out of court the possibility that ordinary objects are metaphysically emergent.

I start by considering whether any ordinary objects are either Weakly emergent or (as I will sometimes put it) are 'at least' Weakly emergent, in having at least one feature satisfying the conditions in the schema for Weak emergence (§6.1). I offer three routes to an affirmative answer. First, I argue that ordinary objects of the sort appropriately treated by classical (or 'Newtonian') me-

chanics are Weakly emergent by lights of a DOF-based account, thanks to the elimination of quantum DOF in the classical limit (§6.1.1); second, I argue that a common conception of artifacts as associated with sortal properties and distinctive functional roles, and the associated compositionally flexible persistence conditions typically encoded in these sortal features, supports thinking of artifacts as being at least Weakly emergent by lights of a functional realization account (§6.1.2); third, I argue that ordinary objects typically have metaphysically indeterminate boundaries, which when coupled with an attractive determinable-based account of such indeterminacy (advanced in my 2013 and 2016a), indicates that such ordinary objects are at least Weakly emergent, by lights of a determinable-based account of such emergence (§6.1.3).

I next consider whether any ordinary objects are Strongly emergent (§6.2). I argue that the best case for this stems from the role mentality plays in both the specification and the constitution of the functional roles (typically encoding social practices involving normative or aesthetic goings-on) which are typically associated with artifacts. The ultimate status of such objects as Strongly or rather just Weakly emergent hinges, like the status of certain complex systems involving mentality, on the status as Weakly or Strongly emergent of the associated mental features of persons, of the sort to be discussed in the next chapters.

I close by observing that the results of this chapter undercut the motivations for Thomasson's meta-ontological view, as discussed in her (2010) and elsewhere, according to which investigations into the ontological status of artifactual ordinary objects should proceed differently from investigations into the ontological status of special-science entities (§6.3). Thomasson's suggestion is primarily motivated by thinking, first, that the usually stated concerns with ordinary objects (e.g., Kim-style causal overdetermination concerns) arise from trying to give scientific and ordinary objects (including artifacts) a unified treatment, and second, that the concerns as attaching to scientific goings-on do not admit of any good answers. But as I have argued, there are good responses to the concerns at issue, whether natural or artifactual ordinary objects are at issue. Nothing stands in the way of a systematic treatment of natural and artifactual ordinary objects as at least Weakly emergent, and—contingent upon future empirical results and the import of mentality to be next considered—perhaps even Strongly emergent.

Chapter 7: "Consciousness"

In Chapter 7, I turn to considering whether consciousness or conscious experience of the sort that we and other creatures enjoy is either Weakly or Strongly emergent. There are many forms or species of consciousness, including perceptual awareness of the external world, conscious awareness of internal states (e.g., pain), and self-consciousness (i.e., consciousness of ourselves as conscious beings). Little in this chapter hinges on differences between these forms of consciousness, so I speak generically of consciousness or conscious awareness (or associated mental features), which may have as its seeming object the external world, one's internal states, or (as a special case of the latter) consciousness itself.

I start by considering whether consciousness is Strongly emergent (§7.1). Arguments for consciousness's being Strongly emergent (or in any case physically unacceptable, in a way compatible with being Strongly emergent) typically

rest on the commonly accepted failure of consciousness to be predictable from or explainable in terms of lower-level physical phenomena. Although for reasons mentioned previously, even in-principle epistemic failures can't be the whole story, proponents of these arguments offer reasons for thinking that the explanatory gaps are taken to be metaphysically significant, in reflecting not just mathematical barriers to explanation (e.g., non-linearity), but rather that the subjective or qualitative aspects of conscious experience depart so greatly from lower-level physical features that no physicalist account of consciousness can be correct. I consider the two most promising forms of explanatory gap argument, however, and argue that neither goes through.

I first address knowledge arguments (per Nagel 1974 and Jackson 1982 and 1986) aiming to show that one could have complete physical knowledge of some entity or subject matter, but nonetheless fail to know certain facts pertaining to conscious states associated with the entity or subject matter (§7.1.1). I focus on Jackson's case-based argument, whereby Mary, a scientist confined to a black and white room, comes to possess complete physical knowledge about human color vision; but upon being released and seeing a ripe tomato, learns something new—such that, the conclusion goes, physicalism is thereby revealed to be false. Much physicalist ink has been spilled on responding to Jackson's argument; here I advance a response not much on the books, which proceeds by denying that Mary has complete physical knowledge about human color vision before her release, per what I call the 'Incomplete Physical Knowledge' strategy. I motivate this strategy by observing that a physicalist need not agree that physical knowledge must be 'objective' in the sense of failing to be of subjective or qualitative aspects of reality, since such a view is in tension with physicalism—which maintains, after all, that some sufficiently complex physical goings-on are identical with or realize conscious mental states and associated subjective/qualitative features. Relatedly, I maintain, the physicalist can and arguably should simply grant that acquaintance is a necessary condition for knowing certain physical facts—namely, those providing a constitutive basis for any subjective or qualitative aspects of consciousness there may be. I note certain advantages that the Incomplete Physical Knowledge strategy has over other responses, and diagnose the failure for this strategy to be properly appreciated as reflecting a mistaken characterization of the physical goings-on in overly representational, insufficiently expansive (i.e., appropriately complex), and qualitatively etiolated terms. The upshot is that the knowledge arguments do not provide compelling reason to think that consciousness and its associated subjective and qualitative aspects are actually physically unacceptable, much less actually Strongly emergent.

I next address the conceivability argument advanced and developed by Chalmers (in his 1996, 1999, 2009, and elsewhere), according to which the conceivability of zombies—creatures which are functional and physical duplicates of creatures like us, but which are lacking in any conscious mentality—is taken, in combination with certain other commitments, to establish the Strong emergence of consciousness (§7.1.2). Chalmers's argument goes beyond previous explanatory gap arguments in that the conceivability of zombies is situated in an independently motivated framework—'epistemic two-dimensionalism' (E2D)—according to which certain facts about meaning, which are taken to be a priori accessible, can be used to identify or establish certain facts about modality, expressing or encoding what is genuinely metaphysically possible (necessary, contingent, impossible). It is commonly assumed that the mode of a priori access to

meanings that enters into the E2D strategy proceeds by way of conceiving. Consequently, commitment to the E2D strategy for gaining (much) access to modal truth, and to implementing this strategy via a conceiving-based epistemology of meanings, provides an independent basis for taking the conceivability of zombies to have anti-physicalist metaphysical import, as reflecting a systematic connection between conceivability and metaphysical possibility. The conceivability argument then proceeds as follows:

1. It is conceivable that there is a world which is physically exactly like our world, but in which there is no consciousness.
2. If the world described in (1) is conceivable, then it is metaphysically possible. (E2D)
3. If the world described in (1) is metaphysically possible, then physicalism is false.
4. Physicalism is false.
5. In particular, consciousness is physically unacceptable (and moreover might be Strongly emergent).

The focus of my critical attention here is on the second premise. Drawing on Biggs and Wilson 2017*a* and 2019, I suggest that there is an alternative, and superior, way in which the E2D strategy might be implemented—namely, by appeal to an abduction-based rather a conceiving-based epistemology of the meanings entering into this strategy. I then argue that it is far from clear that the genuine possibility of zombies, or the associated Strong emergence of consciousness, is output from E2D, when this framework is implemented using abduction rather than conceiving. One might wonder, as against this line of thought, whether abduction is apt for purposes of implementing E2D, given that (as above) the access to the meanings which are in turn supposed to provide a basis for access to modal truths is supposed to proceed in a priori fashion. Here again, I draw on joint work with Biggs (Biggs and Wilson 2017*b*), where we argue that, contra common assumption, abduction is an a priori mode of inference—as a priori as conceiving, in particular.[11] The upshot is that, like the knowledge arguments, Chalmers's two-dimensional argument fails to establish that consciousness is actually physically unacceptable, much less Strongly emergent.

I go on to consider whether consciousness is Weakly emergent (§7.2). Here I argue for an affirmative answer, based in the fact that qualitative conscious states—e.g., states of conscious awareness of colors or pains—are typically determinable rather than (maximally) determinate, in a way that defensibly renders them suitable (again, assuming that they are not Strongly emergent) for being realized in determinable-based fashion, and hence Weakly emergent. I first provide two reasons for thinking that various of our perceptions are determinable (§7.2.1), the first being that qualitative mental states are susceptible to Sorites phenomena, and the second reflecting that our perception of macro-entities and

[11] Such a view is not as unusual as it might first appear. To start, the view has precursors in Kant (via the notion of the synthetic a priori) and Carnap (and his appeal to conceptual analysis as involving 'explication,' which proceeds abductively). Moreover, the view reflects the underappreciated fact that the ceteris paribus clauses in abductive principles (e.g., one or other principle of parsimony) effectively operate to shield them from disconfirmation. See our papers for further details.

their features typically fails to register micro-determinate details. Now, as previously, one implementation of the schema for Weak emergence is a determinable-based account of realization, according to which it suffices for the realization of a feature that the feature be a determinable of lower-level physical determinates. So, if the determinable qualitative conscious states at issue can be seen as having lower-level physical determinates, we will be in position to conclude that such conscious features are Weakly emergent.

I then present arguments, due to Ehring (1996), Funkhouser (2006), and Walter (2006), according to which this does not make sense; here the common line is that while the determinable/determinate relation has some feature *F*, the relation between qualitative conscious states and lower-level physical states does not have *F* (§7.2.2). For example, Ehring argues that taking qualitative conscious features to be determinables of lower-level physical determinates is incompatible with the intuitive possibility of there being qualitative mental superdeterminates (e.g., a maximally specific pain), since implying, falsely, that these could be further determined. Drawing on my (2009), I respond to Ehring's and the other concerns by noting, first, that different sciences may treat a single determinable as having different determination dimensions (hence mental features may be superdeterminate relative to a purely psychological science, while being further determined relative to a lower-level physical science), and second, arguing that a proper understanding of the determinable/determinate relation, per

> *Powers-based Determination*: feature P is a determinate of feature Q iff Q is associated with a proper subset of the powers associated with P, and the set of powers had by P but not by Q is not associated with any property,

provides a comprehensible metaphysical basis for accommodating the phenomenon of science-relative determination dimensions. To wit: relative to one set of determination dimensions, reflecting sensitivity to powers associated with the determinable set, a given qualitative conscious state might be characterized as a superdeterminate; but relative to a finer-grained set of determination dimensions (reflecting sensitivity to powers in relevant supersets of the determinable set) that same feature might not be appropriately characterized as a superdeterminate (§ 7.2.3).

Chapter 8: "Free Will"

Free will (or free agency), if such there be, involves the ability to mentally choose an outcome (an intention to ϕ, or a ϕ-ing), where the outcome is 'free' in being, in some substantive sense, up to the agent of the choice. In Chapter 8, I consider whether free will of the sort that we appear to have and to exercise is either Weakly or Strongly emergent.

I start by drawing on Bernstein and Wilson 2016 in order to set up a useful framework for investigating into whether free will is metaphysically emergent (§8.1). Recall that the schemas for Weak and Strong emergence were initially motivated as associated with two specific responses to the problem of higher-level causation. Mental features are a common focus of this problem, but in the usual case the mental features at issue are qualitative or intentional features, for which free choice is supposed not to be at issue. More generally, debates over the status of free will have tended to proceed in relative independence from debates over the status of mental features whose governance by natural law is taken for granted. As Bernstein and I argue, however, the problematics underlying

the free will and the mental causation debates are appropriately seen as special cases of a more general problem, concerning whether and how mental features of a given type may be efficacious, qua the types of feature they are (qualitative, intentional, freely deliberative), given their apparent causal irrelevance—i.e., apparent failure of distinctive efficacy—for effects of the type in question. That the free will and mental causation debates can be seen as special cases of a more general problem serves to suggest certain parallels between positions in the respective debates, which parallels are useful for purposes of assessing whether free will is either Weakly or Strongly emergent.

In the next two sections I develop these parallels for compatibilism and libertarianism, respectively. Again drawing on Bernstein and Wilson 2016, I first argue that a representative range of compatibilist accounts, including accounts of freedom as underdetermination (per, e.g., Ayer 1954), freedom as ownership (per, e.g., Davidson 1963), and freedom as responsibility (per, e.g., Strawson 1962), implement a structurally similar 'proper subset' strategy for responding to the problem of free will (§8.2). Effectively, the general compatibilist strategy is to identify a proper subset of the total causal antecedents of a given outcome (effect) of a mental choosing, as that which is relevant for the choosing's being efficacious qua free; different compatibilists then differ about which proper subsets of the total causal antecedents are those which are so relevant. I then extend this result, arguing that the compatibilist strategy can be more specifically understood as entailing the holding of a proper subset relation between token powers associated with two complex, cotemporal events, corresponding to, first, the mental choosing M in combination with the relevant causal antecedents of M (call this complex event C'), and second, the mental choosing M in combination with the total causal antecedents of M (call this complex event C). I next argue that a representative range of libertarian accounts, including event-causal accounts (per, e.g., Kane 1996 and Merricks 2003), agent-causal accounts (per, e.g., O'Connor 2005), and 'non-causal'[12] accounts (per, e.g., Ginet 1990, McCann 1998, and Stump 1999) are reasonably seen as committed to free will's being associated with a fundamentally novel power—namely, the power to freely choose to ϕ—not had by lower-level physical goings-on, of the sort that satisfaction of the schema for Strong emergence requires (§8.3).

Parallels established, I turn to considering whether (some cases of) free will might be Weakly emergent (§8.4.1). The prospects are good, I argue. Though free choices are not taken to be part of a higher-level system of laws on either compatibilist or libertarian accounts, a compatibilist account is one manifesting the usual Weak emergentist characterization of special-science goings-on as comparatively insensitive to lower-level physical details, in the sense that an agent's reasons for action in a given case float free of many such details (and in particular, are sensitive only to facts about 'relevant' causal antecedents). Since our deliberations and associated acts of choice clearly are insensitive to many microphysical details, then given that free will is understood along compatibilist (Weak emergentist) lines, there is good reason to think that such free will actually exists, and moreover is abundant.

[12] Note that non-causal accounts of libertarian free will only require that the choice not be antecedently caused; they are compatible with, and indeed require, that the choice itself be efficacious (hence have powers).

Notwithstanding that there is presumably plenty of what compatibilists count as free will, is there actually free will of a libertarian, nomologically transcendent variety (§8.4.2)? I offer a new argument for an affirmative answer, as follows:

1. We experience ourselves as seeming to freely choose, in ways transcending any nomological (deterministic or indeterministic) goings-on.
2. In the absence of good reasons to think that our experience of nomologically transcendent free will cannot be taken at face value, we are entitled to take this experience at face value.
3. There are no good reasons to think that our experience cannot be taken at face value.
4. We are entitled to take our experience of nomologically transcendent free will at face value.

The argument is valid, and premise (1) is clearly true (even non-libertarians agree). Premise (2) also seems reasonable: if we have clear experience of some seeming phenomenon, we need good reason not to take that experience at face value. I focus on defending premise (3) against the 'Libet cases' which pose the most serious challenge to taking our experience at face value.

Recall that Libet (1999) determined that when a subject is asked to move their finger and track exactly when the urge to do so occurs, an unconscious 'Readiness Potential' RP precedes the "experience of will" by around 400 milliseconds. Libet and others concluded that conscious will is not the initiator of voluntary action, but instead a consequence of an unconscious physical process that triggers the action. In response, I first canvass certain alternative interpretations of the data, due to Mele (2009) and O'Connor (2005), which are compatible with nomologically transcendent free will. I then offer a new interpretation of my own, which is also so compatible, and which takes advantage of the cotemporal material dependence condition in Strong emergence. On my interpretation, the intention to choose and the associated brain activity are cotemporally initiated, but it takes a bit of time for this fact to consciously register as a complete thought in the agent's mind. Thinking takes time—more time, perhaps, than a choice. A very small lag—less than half a second—would be a natural concomitant of our mental decision-making processes, compatible with transcendent free will. Correspondingly, Libet's assumption that "In the traditional view […], one would expect conscious will to appear before, or at the onset, of the RP, and thus command the brain to perform the intended act" (1999, 49) reflects an overly simplistic account of how nomologically transcendent free will would actually work.

Chapter 9: "Closing Remarks"

In Chapter 9, I summarize the results of the book and call attention to some phenomena whose status as metaphysically emergent deserves further attention, including quantum entanglement, molecular structure, biological systems, brain dynamics, and spacetime. I close with some methodological observations pointing towards other ways in which attention to broadly mereological relationships between sets of powers might serve to shed light on other aspects of higher-level reality, beyond metaphysical emergence.

References

Aizawa, K. and Gillett, C., 2009. The (multiple) realization of psychological and other properties in the sciences. *Mind and language*, 24, 181–208.

Alexander, S., 1920. *Space, time, and deity*. London: Macmillan.

Antony, L.M. and Levine, J.M., 1997. Reduction with autonomy. *Philosophical perspectives*, 11, 83–105.

Ayer, A.J., 1954. Freedom and necessity. *In*: A.J. Ayer, *Philosophical Essays*. London: Macmillan, 271–284.

Baltimore, J.A. 2013. Careful, physicalists: Mind–body supervenience can be too superduper. *Theoria*, 79, 8–21.

Barnes, E., 2012. Emergence and fundamentality. *Mind*, 121, 873–901.

Batterman, R.W., 1998. Why equilibrium statistical mechanics works: Universality and the renormalization group. *Philosophy of science*, 65, 183–208.

Batterman, R.W., 2000. Multiple realizability and universality. B*ritish journal for the philosophy of science*, 51, 115–45.

Batterman, R.W., 2002. *The devil in the details: Asymptotic reasoning in explanation, reduction, and emergence*. Oxford: Oxford University Press.

Baysan, U. and Wilson, J., 2017. Must strong emergence collapse? *Philosophica*, 91, 49–104.

Bedau, M.A., 1997. Weak Emergence. *In*: J. Tomberlin, ed. *Mind, causation and world, philosophical perspectives annual volume 11*. Oxford: Wiley-Blackwell, 375–399.

Bedau, M.A., 2008. Is weak emergence just in the mind? *Minds and machines*, 18, 443–459.

Bennett, K., 2003. Why the exclusion problem seems intractable and how, just maybe, to tract it. *Noûs*, 37, 471–497.

Bennett, K., 2017. *Making things up*. Oxford: Oxford University Press.

Bernstein, S. and Wilson, J.M., 2016. Free will and mental quausation. *Journal of the American philosophical association*, 2, 310–331.

Biggs, S. and Wilson, J.M., 2017a. Abductive two-dimensionalism: A new route to the a priori identification of necessary truths. *Synthese*, 197, 59–93.

Biggs, S. and Wilson, J.M., 2017b. The a priority of abduction. *Philosophical studies*, 174, 735–758.

Biggs, S. and Wilson, J.M., 2019. Abduction vs. conceiving in modal epistemology. *Synthese*, 198, 2045–2076.

Boyd, R., 1980. Materialism without reduction: What physicalism does not entail. *In*: N. Block, ed. *Readings in the philosophy of psychology*. Cambridge (Mass.): Harvard University Press, 67–106.

Broad, C.D., 1925. *Mind and its place in nature*. Cambridge: Kegan Paul.

Carruth, A., 2018. Emergence, reduction and the identity and individuation of powers. *Topoi*. 39, 1021–1030.

Chalmers, D.J., 1996. *The conscious mind*. Oxford: Oxford University Press.

Chalmers, D.J., 1999. Materialism and the metaphysics of modality. *Philosophy and phenomenological research*, 59, 473–96.

Chalmers, D.J., 2006. Strong and weak emergence. *In*: P. Clayton and P. Davies, eds. *The re-emergence of emergence*. Oxford: Oxford University Press, 244–256.

Chalmers, D.J., 2009. The two-dimensional argument against materialism. *In*: B.P. McLaughlin and S. Walter, eds. *Oxford Handbook to the Philosophy of Mind*. Oxford: Clarendon Press.

Clapp, L., 2001. Disjunctive properties: Multiple realizations. *Journal of philosophy,* 98, 111–136.

Cleve, J. van., 1990. Mind-dust or magic? Panpsychism versus emergence. *Philosophical perspectives*, 4, 215–226.

Crane, T., 2001. The significance of emergence. *In*: C. Gillett and B. Loewer, eds. *Physicalism and its discontents*. Cambridge: Cambridge University Press, 207–224.

Craver, C.F., 2001. Role functions, mechanisms, and hierarchy. *Philosophy of science*, 68, 53–74.

Cummins, R. 1975. Functional analysis. *Journal of philosophy*. 72, 741–764.

Cunningham, B., 2001. The reemergence of emergence. *Philosophy of science*, 68, S62–S75.

Dasgupta, S., 2014. The possibility of physicalism. *Journal of philosophy*, 111, 557–592.

Davidson, D., 1963. Actions, reasons, and causes. *Journal of philosophy*, 60, 685–700.

Dosanjh, R., 2014. *A defense of reductive physicalism*. Thesis (PhD). University of Toronto.

Dosanjh, R., 2019. Token-distinctness and the disjunctive strategy. *Erkenntnis*, 86, 715–732.

Douven, I., 2021. Abduction. *In*: E. Zalta, ed. *The Stanford encyclopedia of philosophy*, Stanford University, summer 2021 edition.

Ehring, D., 1996. Mental causation, determinables, and property instances. *Noûs*, 30, 461–480.

Field, H., 2003. Causation in a physical world. *In*: M. Loux and D. Zimmerman, eds. *The oxford handbook of metaphysics*. Oxford: Oxford University Press.

Fodor, J., 1974. Special sciences (or, the disunity of science as a working hypothesis). *Synthese*, 28, 77–115.

Fodor, J., 1987. *Psychosemantics*. Cambridge, MA: MIT Press.

Francescotti, R., 2007. Emergence. *Erkenntnis*, 67, 47–63.

Funkhouser, E., 2006. The determinable-determinate relation. *Noûs*, 40, 548–569.

Gillett, C., 2002. The dimensions of realization: A critique of the standard view. *Analysis*, 62, 316–323.

Gillett, C., 2016. *Reduction and emergence in science and philosophy*. Cambridge: Cambridge University Press.

Ginet, C., 1990. *On action*. Cambridge: Cambridge University Press.

Harman, G.H., 1965. The inference to the best explanation. *Philosophical review*, 74, 88–95.

Haug, M.C., 2010. Realization, determination, and mechanisms. *Philosophical studies*, 150, 313–330.

Heil, J., 2003. Levels of reality. *Ratio*, 16, 205–221.

Hempel, C. and Oppenheim, P., 1948. Studies in the logic of explanation. *Philosophy of science*, 15, 135–175.

Horgan, T., 1993. From supervenience to superdupervenience: Meeting the demands of a material world. *Mind*, 102, 555–586.

Howell, R.J., 2009. Emergentism and supervenience physicalism. *Australasian Journal of philosophy*, 87, 83–98.

Humphreys, P., 1996. Aspects of emergence. *Philosophical topics*, 24, 53–70.

Jackson, F., 1982. Epiphenomenal qualia. *Philosophical quarterly*, 32, 127–136.

Jackson, F., 1986. What Mary didn't know. *The journal of philosophy*, 83, 291–295.

Jaworski, W., 2002. Multiple-realizability, explanation and the disjunctive move. *Philosophical studies*, 108, 289–308.

Kane, R., 1996. *The significance of free will*. New York: Oxford University Press.

Kim, J., 1992. 'Downward causation' in emergentism and nonreductive physicalism'. *In*: A. Beckermann, H. Flohr, and J. Kim, eds. *Emergence or reduction? Prospects for nonreductive physicalism*, Berlin: De Gruyter, 119–138.

Kim, J., 1999. Making sense of emergence. *Philosophical studies*, 95, 3–36.

Kim, J., 2010. Thoughts on Sydney Shoemaker's physical realization. *Philosophical studies*, 148, 101–112.

Klee, R., 1984. Micro-determinism and concepts of emergence. *Philosophy of science,* 51, 44–63.

Ladyman, J. and Ross, D., 2007. *Every thing must go: Metaphysics naturalized*. Oxford: Oxford University Press.

Lamb, M., 2015. *Characteristics of non-reductive explanations in complex dynamical systems research*. Thesis (PhD). University of Cincinnati.

LePore, E. and Loewer, B., 1987. Mind matters. *The journal of philosophy*, 84, 630–642.

LePore, E. and Loewer, B., 1989. More on making mind matter. *Philosophical topics*, 17, 175–191.

Lewes, G.H., 1875. *Problems of life and mind*. London: Kegan Paul, Trench, Turbner & Co.

Libet, B.W., 1999. Do we have free will? *Journal of consciousness studies*, 6 (8-9), 47–57.

MacDonald, C. and MacDonald, G., 1986. Mental causes and explanation of action. *In*: L. Stevenson, R. Squires, and J. Haldane, eds. *Mind, causation, and action*. Oxford: Basil Blackwell, 145–158.

McCann, H., 1998. *The works of agency: On human action, will, and freedom*. Ithaca: Cornell University Press.

McLaughlin, B., 1992. The rise and fall of British emergentism. *In*: A. Beckerman, H. Flohr, and J. Kim, eds. *Emergence or reduction? Essays on the prospects of nonreductive physicalism*. Berlin: De Gruyter, 49–93.

McLaughlin, B.P., 2007. Mental causation and Shoemaker-realization. *Erkenntnis.* 67, 149–172.

Megill, J., 2013. A defense of emergence. *Axiomathes*, 23, 597–615.

Mele, A. 2009. *Effective intentions: The power of conscious will*. Oxford: Oxford University Press.

Melnyk, A., 2003. *A physicalist manifesto: Thoroughly modern materialism*. New York: Cambridge University Press.

Melnyk, A., 2006. Realization-based formulations of physicalism. *Philosophical studies*, 131, 127–155.

Merricks, T. 2003. *Objects and persons*. Oxford: Clarendon Press.

Mill, J.S., [1843]1973. *A system of logic*. Toronto: University of Toronto Press.

Mitchell, S.D. 2012. Emergence: Logical, functional and dynamical. *Synthese*, 185, 171–186.

Morgan, C.L., 1923. *Emergent evolution*. London: Williams & Norgate.

Morris, K., 2018. *Physicalism deconstructed: Levels of reality and the mind–body problem*. Cambridge: Cambridge University Press.

Morrison, M., 2012. Emergent physics and micro-ontology. *Philosophy of science*, 79, 141–166.

Nagel, E., 1961. *The structure of science*. London: Routledge & Kegan Paul.

Nagel, T., 1974. What is it like to be a bat? *The philosophical review*, 83, 435–450.

Newman, D., 1996. Emergence and strange attractors. *Philosophy of science*, 63, 245–261.

Ney, A., 2010. Convergence on the problem of mental causation: Shoemaker's strategy for (nonreductive?) physicalists. *Philosophical issues*, 20, 438–445.

Ney, A., 2022. Review of *Metaphysical emergence* by Jessica Wilson. *Notre Dame philosophical reviews.*

Noordhof, P., 2010. Emergent causation and property causation. *In*: C. Macdonald and G. Macdonald, eds. *Emergence in mind*. Oxford: Oxford University Press, 69–99.

O'Connor, T., 1994. Emergent properties. *American philosophical quarterly*, 31, 91–104.

O'Connor, T., 2005. Free will. *In*: J.M. Fischer, ed. *Free will: Critical concepts in philosophy*. New York: Routledge, 7–32.

O'Connor, T. and Wong, H.Y., 2005. The metaphysics of emergence. *Noûs*, 39, 658–678.

Paolini Paoletti, M., 2017. *The quest for emergence*. Munich: Philosophia.

Papineau, D., 1993. *Philosophical naturalism*. Oxford: Basil Blackwell.

Pereboom, D., 2002. Robust non-reductive materialism. *Journal of philosophy*, 99, 499–531.

Poland, J., 1994. *Physicalism: The philosophical foundations*. Oxford: Clarendon Press.

Polger, T.W. 2007. Realization and the metaphysics of mind. *Australasian journal of philosophy*, 85, 233–259.

Putnam, H., 1967. Psychological predicates. *In*: D.D. Merrill and W. Capitan, eds. *Art, mind, and religion*. Pittsburgh: University of Pittsburgh Press, 37–48.

Rueger, A. and McGivern, P., 2010. Hierarchies and levels of reality. *Synthese*, 176, 379–397.

Russell, B., 1912. On the notion of cause. *Proceedings of the Aristotelian society*, 13, 1–26.

Schaffer, J., 2009. On what grounds what. *In*: D. Manley, D. Chalmers, and R. Wasserman, eds. *Metametaphysics: New essays on the foundations of ontology*, Oxford: Oxford University Press, 347–383.

Schroder, J., 1998. Emergence: non-deducibility or downward causation? *The philosophical quarterly*, 48, 433–452.

Seager, W., 1999/2016. *Theories of consciousness: an introduction and assessment*. London: Routledge.

Searle, J.R., 1992. *The rediscovery of the mind*. Cambridge (Mass.): MIT Press.

Shoemaker, S., 2000/2001. Realization and mental causation. *In*: Eds. *Proceedings of the 20th world congress in philosophy*. Cambridge: Philosophy Documentation Center, 23–33.

Silberstein, M. and McGeever, J., 1999. The search for ontological emergence. *Philosophical quarterly*, 50, 182–200.

Sperry, R., 1986. Discussion: macro- versus micro-determination. *Philosophy of science*, 53, 265–270.

Stephan, A., 2002. Emergentism, irreducibility, and downward causation. *Grazer philosophische studien*, 65, 77–93.

Strawson, P.F., 1962. Freedom and resentment. *Proceedings of the British academy*, 48, 1–25.

Stump, E., 1999. Dust, determinism, and Frankfurt. *Faith and philosophy*, 16, 413–422.

Tahko, T.E., 2018. Fundamentality. *In*: E. Zalta, ed. *The Stanford encyclopedia of philosophy*. Stanford University, Fall 2020 edition.

Taylor, E., 2015. Collapsing emergence. *Philosophical quarterly*, 65, 732–753.

Thomasson, A.L., 2007. *Ordinary objects*. Oxford: Oxford University Press.

Thomasson, A.L., 2010. The controversy over the existence of ordinary objects. *Philosophy compass*, 5, 591–601.

Thompson, E. and Varela, F.J., 2001. Radical embodiment: neural dynamics and consciousness. *Trends in cognitive sciences*, 5, 418–425.

Van Gulick, R., 2001. Reduction, emergence and other recent options on the mind/body problem: a philosophic overview. *Synthese*, 8, 1–34.

Walter, S., 2006. Determinates, determinables, and causal relevance. T*he Canadian journal of philosophy*, 37, 217–243.

Walter, S., 2010. Taking realization seriously: no cure for epiphobia. *Philosophical studies*, 151, 207–226.

Wilson, J.M., 1999. How superduper does a physicalist supervenience need to be?' *The philosophical quarterly*, 49, 33–52.

Wilson, J.M., 2002. Causal powers, forces, and superdupervenience. *Grazer philosophische-studien*, 63, 53–78.

Wilson, J.M., 2005. Supervenience-based formulations of physicalism. *Noûs,* 39, 426–459.

Wilson, J.M., 2006. On characterizing the physical. *Philosophical studies*, 131, 61–99.

Wilson, J.M., 2009. Determination, realization, and mental causation. *Philosophical studies*, 145, 149–169.

Wilson, J.M., 2010. Non-reductive physicalism and degrees of freedom. *British journal for the philosophy of science*, 61, 279–311.

Wilson, J.M., 2011. Much ado about 'something': Critical notice of Chalmers, Manley, Wasserman, *Metametaphysics*. *Analysis*, 71, 172–188.

Wilson, J.M., 2013. A determinable-based account of metaphysical indeterminacy. *Inquiry*, 56, 359–385.

Wilson, J.M., 2014. No work for a theory of grounding. *Inquiry*, 57, 1–45.

Wilson, J.M., 2015. Metaphysical emergence: Weak and strong. *In*: T. Bigaj and C. Wüthrich, eds. *Metaphysical emergence in contemporary physics; Poznan studies in the philosophy of the sciences and the humanities*, Amsterdam/New York: Brill, 251–306.

Wilson, J.M., 2016a. Are there indeterminate states of affairs? Yes. *In*: E. Barnes, ed. *Current controversy in metaphysics,* London: Routledge, 105–119.

Wilson, J.M., 2016b. The question of metaphysics. *The philosophers' magazine*, 74, 90–96.

Wilson, J.M., 2016c. Three barriers to philosophical progress. *In*: D. Broderick and R., Blackford, eds. *Philosophy's future: the problem of philosophical progress*, Oxford: Wiley-Blackwell, 91–104.

Wilson, J.M., 2021. *Metaphysical emergence*. Oxford: Oxford University Press.

Wilson, J.M., 2022/2017. Determinables and determinates. In E. Zalta, ed. *The Stanford encyclopedia of philosophy*. Stanford University, Spring 2023 edition.

Wilson, J.M., forthcoming*a*. The fundamentality first approach to metaphysical structure. A*ustralasian journal of philosophy*.

Wilson, J.M., forthcoming*b*. On the notion of diachronic emergence. In: D. Yates, ed. *Rethinking emergence*.

Wilson, J.M., under contract. *Fundamentality and metaphysical dependence*. Oxford: Oxford University Press.

Wimsatt, W., 1996. Aggregativity: Reductive heuristics for finding emergence. *Philosophy of science*, 64, 372–384.

Witmer, D.G., 2003. Functionalism and causal exclusion'. *Pacific philosophical quarterly*, 84, 198–214.

Yablo, S., 1992. Mental causation. *The philosophical review*, 101, 245–280.

Yates, D., 2012. Functionalism and the metaphysics of causal exclusion. *Philosophers' imprint*. 12, 1–25.

Yates, D., 2016. Demystifying emergence. *Ergo: an open access journal of philosophy*, 3, 809–841.

Biochemical Functions as Weakly Emergent

Francesca Bellazzi

University of Birmingham

Abstract

This paper will consider how the account of weak emergence presented by Wilson in the book *Metaphysical emergence* (2021) can be used to explore the relation between biochemical functions and chemical structure in biochemical molecules, as vitamin B12. The structure of the paper is the following. Section 2 will introduce why biochemical functions are interesting from a philosophical perspective and why their relation to molecular structure can be seen as problematic. In doing so, it will consider the definition of biochemical functions as in Bellazzi (2022) for which they can be seen as sets of chemical dispositional properties that contribute to biological processes. Section 3 will explore how, given this definition of biochemical functions, we can interpret the relation between chemical structure and biochemical structure via weak emergence. Section 4 concludes.

Keywords: Metaphysical emergence, Biochemical functions, Biochemical kinds.

1. Introduction

This paper will consider how the account of weak emergence presented by Wilson in the book *Metaphysical Emergence* (2021) can be used to explore the relation between biochemical functions and chemical structure in biochemical molecules, as vitamin B12. The discussion of the relation between biochemical function and chemical structure is relevant to the debate concerning inter-level relations together with being a foundational topic for biochemistry (Santos et al. 2020).[1] Moreover, the results of this paper provide a novel application of Wilson's account of weak emergence, enriching the case studies that can fit with the framework and offering new insights into the understanding of weak emergence in non-yet-considered cases.

The structure of the paper is the following. Section 2 will introduce why biochemical functions are interesting from a philosophical perspective and why their relation to molecular structure can be seen as problematic. In doing so, it will

[1] This paper draws on some of the results in Bellazzi 2023.

Argumenta 10, 1 (2024): 225—235
ISSN 2465-2334
DOI 10.14275/2465-2334/202418.bel
First published: 30 November 2024

consider the definition of biochemical functions as in Bellazzi 2022 for which they can be seen as sets of chemical dispositional properties that contribute to biological processes. Section 3 will explore how, given this definition of biochemical functions, we can interpret the relation between chemical structure and biochemical structure via weak emergence. Section 4 concludes.

2. Structure and Function in Biochemical Kinds

Chemistry is often taken to be the domain of chemical structure and kinds characterized in micro-structural terms, such as constituent atomic properties.[2] Biology, instead, is the domain of evolutionary functions, etiological classifications and pluralism (Slater 2009; Bartol 2016). Biochemistry stands as an hybrid domain between the two. While it is not easy to provide a set of necessary and sufficient conditions for a kind to be biochemical, the literature on the topic agrees that biochemical kinds need to exhibit at least two kinds of properties: structural ones and functional ones (Slater 2009; Bartol 2016; Havstad 2016, 2018; Kistler 2018; Tahko 2020). Proteins, for example, are characterised in terms of structure, the amino-acid chain that composes them, and in terms of the functional roles that they play within biological systems.

Prima facie, this definition or the combination of these two sets of properties might not be particularly problematic, however the exact relation between structural and functional properties still posits questions (Bartol 2016; Tahko 2020). One of the reasons why this is so is based on the complexity of the relations between structure and function, as they often take the form of multiple realisability and multiple determinability. Multiple realisability (MR) refers to a phenomenon in which the same entity or property can be realised by different ones.[3] For example, the property of being an eye can be realised by different organs in different animals. Multiple determinability (MD) refers to the opposite phenomenon: when the same entity can determine different properties or other entities. For example, the same chemical compound can enter into different chemical reactions, realising different properties.

In the biochemical case, MR and MD are particularly relevant because the same biochemical function can be realised by multiple microstructures and the same microstructure can realise multiple biochemical functions (Tahko 2020). Two relevant examples in this regard are haemoglobin for MR and the crystalline proteins for MD. As discussed and presented by Tahko (2020, 2021), haemoglobin is a protein with the function of binding and releasing oxygen and can be constituted by at least two different polypeptide chains (or more). The biochemical function of haemoglobin can be considered an instance of MR, as the function of binding and releasing oxygen is realised by at least two distinct macromolecules (chains of polypeptides) that present some micro-structural differences. This can challenge the identification of an identity reductive relation between the chemical structural properties and the functional ones. Multifunctional proteins or "moonlighting" proteins, such as crystallines, represent instances of MD instead. Crystallines are structural proteins present in all vertebrates' eye lenses, having a function in allowing sight, but they can also have an enzymatic role in digestive

[2] Even if this has been challenged as in Tobin 2010, Havstad 2016, 2018.

[3] Realization can be defined as a "synchronic ontological dependence relation, distinct from identity, and that transmits physical legitimacy from physical realizers to what is realized" (Polger and Shapiro 2016: II, 4).

processes. In these cases, we notice a form of MD, as the same chemical structure can lead to very different functions in sight and digestion mechanisms (Tobin 2010; Bartol 2016; Tahko 2020). This again challenges a direct identification of the relation between structure and function, as a strict identity relation between the some underlying structural properties and functional properties does not hold. Moreover, both MR and MD generate issues of taxonomy or classification. If we follow a micro-structuralist approach, then we should favour structure over function and have either many kinds that have the same function (in the case of MR) or one unique kind that has different functions (in the case of MD). If we follow a functional approach, then we have two or three—or as many as the functions—different kinds (in the case of MD) or one kind (in the case of MR).

As a reaction to these tensions, Bartol argues that we should bite the bullet and simply embrace the duality of the two sets of properties: there are chemical structural ones and the biological functional ones (2016). However this approach does not really do justice to the features of biochemical macromolecules that display *both* chemical structure and biological function. These two features are strongly entangled, as supported by some more complex relations between the functions and the chemical structure (see also Goodwin 2011). For instance, Tahko suggests that some cases of MD can be explained or derived from the amphoteric nature[4] of some microstructures (2020). In the cases of some moonlighting proteins for instance, their dual-functions nature can be seen as rooted in some chemical properties of the molecule (Goodwin 2011; Tahko 2020), or at least this can be an option to be analysed in detail.[5] The scientific successes of biochemistry in predicting, manipulating and explaining phenomena encourages instead the exploration of the relation between structure and function, despite its complexity. This is so because this discipline combines chemical and physical model systems to explain and predict biological phenomena.[6]

3. The Double Problem of Biochemical Functions

In order to explore the relation between the chemical structure and biochemical functions one should clarify what are the terms under discussion. Chemical structure comprises both the characterisation of the electronic structure and the molecular geometry of the molecule. What about functional properties? Functional properties in the biochemical context generate what we can call the double problem of biochemical function: the "relation problem" and the "function problem". The "relation problem" asks about the relationship between the chemical structure and the function of a biochemical molecule: how a chemical structure can realise a given biochemical function. As briefly introduced in the previous section, the relation problem is generated by the fact that functional properties in the

[4] An amphoteric chemical substance is one that can react both as a base or as an acid.

[5] The reducibility of the dual nature of moonlighting proteins has been challenged by Santos et al. (2020). This article stresses the importance of analysing the "dynamical interplay between the micro-level of the parts and the macro-level of the relational structures of their systems" in order to understand these proteins (2020: 1). Here I am not supporting the reducibility of biochemical functions to chemical structural properties but rather the relation between functional and structural properties.

[6] The Biochemical Society defines biochemistry as "the branch of science that explores the chemical processes within and related to living organisms" (https://biochemistry.org/education/careers/becoming-a-bioscientist/what-is-biochemistry/).

biochemical domain are often multiply realised, and because biochemical molecules manifest multiple determinability (see Slater 2009; Bartol 2016; Tahko 2020). Furthermore, it is difficult to understand which of the two components, the functional or the structural, has ontological priority in the taxonomy and identification of the biochemical kinds (Slater 2009; Bartol 2016; Tahko 2020). The "function problem" instead asks what biochemical functions are and how they relate to biological functions and the biological component of the kind (Tahko 2020, Bellazzi 2022). Let us consider these problems in more detail with the main case study of this paper, vitamin B12 (as in Bellazzi 2022).

3.1 Vitamin B12

Vitamins B12 are cobalamin chemical compounds that can act as coen-zymes in specific biological processes—specifically, propionate metabolism and methionine biosynthesis. This vitamin comes in four forms—or vitamers—that display similar but different chemical structures: cyanocobalamin, methylcobalamin, hydroxocobalamin, adenosylcobalamin (Combs 2012: 377; Fang et al. 2017).[7] They share a cobalt-corrin complex and the coenzyme function in humans for various biochemical processes such as hematopoiesis, DNA and RNA production, neural metabolism, and carbohydrate, fat, and protein metabolism.[8] Accordingly, these chemical compounds are classified under the same category, 'B12 vitamin', because they display a combination of stable microstructure, a cobalt-corrin complex, and physiological functions.

Vitamin B12 represents an interesting case study relevant to discussing the relation between structure and function because it displays both MR and MD. First, it presents a form of MR in that the biochemical functions of vitamin B12 can be realised by each of the four vitamers recognised in scientific practice.[9] Second, vitamin B12 plays various roles in human physiology, acting in different biological processes, from DNA and RNA production to hematopoiesis, displaying a form of MD too. The combination of MR and MD challenges the identification of simple relations between structure and function. For instance, it makes forms of identity-based reduction, in which the functions of vitamin B12 would be identical to some of the properties of the microstructure, difficult to hold (Tahko 2020). For the sake of the example, let me focus on the function "being a coenzyme in hematopoiesis (the production of blood cells)" (**Coenz-Blood**). B12 vitamers have a biochemical function in the proliferation of erythroblasts (red blood cells) during their differentiation (Koury and Ponka 2004). This happens because vitamin B12 acts as a coenzyme in the reaction involved in regenerating methionine, which is required in normal erythropoiesis. This function is a definitionally important part of the four vitamers of B12: it distinguishes generic cobalt-corrin

[7] A more detailed description is the following: vitamin B12 is "the generic descriptor for all corrinoids (compounds containing the cobalt-centered corrin nucleus) exhibiting qualitatively the biological activity of cyanocobalamin".

[8] Reference for chemical structure and function of vitamin B12 (https://pubchem.ncbi.nlm.nih.gov/compound/Cobalamin). Also, Chapter 12 "B12 Vitamin" in Combs' *The Vitamins: Fundamental Aspects in Nutrition and Health* (2012).

[9] This might represent an instance of multiple constitution of the kind B12, where this kind can be constituted by different chemical compounds that share some functional properties (Kistler 2018). In Kistler, a kind is multiply constituted when it can be constituted by two or more microscopic structures (2018: 18). See also Gillet 2013.

complexes from B12 vitamers, and this shows that, even if it might not be necessary and sufficient on its own to define B12, the functional component is nevertheless important.

Let us go back to the double problem of biochemical functions and elucidate them with the example. First, the "relation problem": **Coenz-Blood** is realised in four different ways via the four vitamers of vitamin B12 and, as such, the relation between the chemical properties of the vitamin B12 and one of its functions should be further explored. The MR of **Coenz-Blood** means that it is at least challenging or not straightforward to map a 1:1 correspondence between it and the possible underlying physicochemical properties. The realisation of this function should be further explored. Second, the "problem of function": what does it mean that vitamin B12 has **Coenz-Blood**} as a *biochemical function*?

The combination of these two problems of biochemical functions might support the suggestion that structure and function could be considered independently. The realisation problem challenges the unification or reduction between the biochemical functions of B12 and its chemical structure. The function problem supports a separation between the chemical and the biological component of biochemical kinds because the nature of biochemical functions could be subsumed under some biological characteristics, which do not relate straightforwardly to the chemical. However, the successes of biochemistry itself seem to provide reasons for the opposite: if we can explain, predict and manipulate biochemical kinds in terms of their function and composition, the two aspects need to be related and, to some extent, ontologically unified.

In order to do so, we should, first, offer a definition of biochemical functions that considers the relation between chemical powers and properties and being dependent on biological context. In this regard, the analysis will start from the following characterisation of biochemical functions (as in Bellazzi 2022):

BC-function: Biochemical functions are associated with a set of chemical powers to bring out a specific effect within biological processes. These biological processes are a product of evolution and, as such, the relevant chemical powers are indirectly evolutionary selected [Fig. 1].

This account of biochemical functions is in line with the general characterisation of biochemistry as the science that considers the behaviour and effects of chemical processes in biological systems (Santos et al. 2020). Moreover, this approach to biochemical functions allows us to answer the function problem, telling us what these properties are, while maintaining the autonomy of the two properties. This provides a starting point to explore the relation between structure and function.

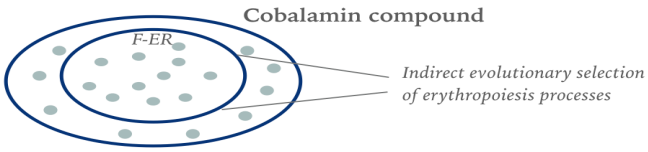


Fig. 1 – The evolutionary selection of the relevant dispositional properties or chemical powers for biochemical functions. In the example, F-ER is **Coenz-Blood** as the function to contribute to erythropoiesis for vitamin B12 as the relevant cobalamin compound.

4. Biochemical Functions as Weakly Emergent

As mentioned in the previous sections, a straightforward form of identity reduction is challenged by the widespread cases of MD and MR in the biochemical domain. Moreover, the set of dispositions relevant to biochemical functions are not any arbitrary chemical powers of the considered molecule or compound but some very specific ones. The relevant powers are those contributing to biological processes and have undergone at least an indirect selection process. The consideration of the biological process they contribute to and—indirectly—evolution that has selected such specific chemical powers is necessary to understand the relevant set of powers (Santos et al. 2020; Bellazzi 2022). Moreover, the causal efficacy of biochemical molecules is distinctive in that it should bring about specific effects within biological processes. Accordingly, an answer to the relation problem should take into account the specificity of biochemical functions together with the relation with structure. In order to provide such an answer, I will consider weak emergence via the proper subset strategy, as in Wilson (2011, 2015, 2021) and as suggested by Tahko (2020). This account, I will suggest, provides an answer to the relation problem and allows for the specificity of biochemical functions.

4.1. Weak Emergence and the "Proper Subset of Powers Strategy"

Weak emergence is a form of emergence compatible with non-reductive physicalism: there is only one broader kind of properties, physical properties. According to non-reductive physicalism, higher-level entities are real and constitute a novel level of reality, being distinctively causally efficacious; at the same time, their causal actions operate in a way respecting physical causal closure and hence in line with physicalism.[10] This combination of distinctiveness and causal efficacy, together with a sense of dependence, can be maintained by defending a form of weak emergence based on the "Proper Subset of Powers strategy" (Wilson 2011, 2021; Tahko 2020).[11] This strategy comprises two steps: i) accepting the *Token Identity of Powers Condition*; ii) accepting the *Proper Subset of Powers Condition*. The first states that every token power of a given token feature H on an occasion *t* is identical with a token power of the token feature L on which H co-temporally materially depends at *t*.[12] The second states that the token feature H has at *t* a non-empty proper subset of the token powers of the token feature L on which H co-temporally materially depends on at *t* (as formulated in Wilson 2021, 57-58). The combination of these two conditions constitutes the basis for a weak emergence relation between the higher and the lower-level entities or features:

[10] The principle of causal closure is often taken as a condition for forms of physicalism and claims that "all physical effects have sufficient physical causes", avoiding cases of problematic overdetermination.

[11] This strategy presupposes a very simple ontology of objects, properties, and powers. Properties are instantiated by objects and are identified by a range of causal powers (Shapiro 2020). In this case, a biochemical molecule instantiates the property "having a given biochemical function", individuated by a specific set of causal powers.

[12] Material dependence implies a form of substance monism, in line with physicalism, and a form of minimal nomological supervenience of the emergent features *type* H on the base features *type* L (Wilson 2021: 73). This means that supervenience should happen with at least nomological necessity.

> **WE**: "What is it for token feature H to be **Weakly Metaphysically Emergent** from token feature L on a given occasion is for it to be the case, on that occasion, i) that H co-temporally materially depends on L, and ii) that H has a non-empty proper subset of the token powers had by L" (Wilson 2021: 75; variables modified, emphasis added).

The first condition i) allows for a form of dependence as there is a token identity of the powers associated with the two features; the second condition ii) allows for a form of distinctiveness. In particular, this account allows for a form of relation between the features because the token powers of the realised feature H are nothing more than a subset of the token powers of a realising feature L, and the two features can be unified as the two sets of powers are both physically acceptable and the token powers of both sets are identical (as also in Shapiro 2020). At the same time, H is ontologically autonomous from L because H has a *proper subset* of the token powers of L and by Leibniz's laws and via set-theory principle, a proper subset of token powers is different from its set of token powers. This permits to maintain the type difference between H and L. The proper subset strategy also allows for a form of causal autonomy, as discussed by Wilson (2011, 2021). Specifically, H has a distinctive causal profile compared to L because it possesses a distinctive set of causal powers or distinctive causal profile compared to L. H's causal autonomy is based on the fact that H has a distinctive set of powers compared to the feature from which it emerges. One of the advantages of this account is that it allows for the relation between the higher and the lower level features, but the higher level ones can still be maintained as ontologically autonomous (Wilson 2011).

Moreover, as will be further detailed in 4.3, the proper subset strategy and weak emergence are able to deal with MR and MD. In the case of MR, it can be possible to identify more than one district token power subset of the lower-level L that can be associated with the higher-level feature H. While in the case of MD, the token set of powers of a given lower-level feature L could present different proper subsets of token powers associated with different higher-level emergent feature H. This allows the account to tackle with some of the issues concerning the relation between structure and function.

4.2 Biochemical Functions Are Weakly Emergent

Let us now consider the interface between biochemical functions and chemical properties and the answer to the relation problem in the light of weak emergence. As in the provided definition, a biochemical function is associated with a set of chemical token powers to bring in a given effect within biological processes (Bellazzi 2022). More precisely, the relation between the token powers associated with the biochemical functions and the correspondent chemical powers can be interpreted with the proper subset view. A biochemical function (BF) has in a given *t* a proper subset of token powers of the set of chemical token powers of the chemical molecule. This proper token subset is individuated via the evolutionary history of the biological process to which BF contributes. Accordingly, following the aforementioned account, we can state the weak emergence of the BF:

> **WE_{BF}**: A biochemical function BF weakly emerges from the chemical compound (C) under consideration at a given *t* because: i) BF co-temporally

materially depends on C at *t*; ii) BF has an identifiable and non-empty proper subset of token powers of C at *t*.

At a given *t*, it is possible to identify the biochemical functions as being associated with a proper subset of the chemical powers, with the powers associated with BF being token identical at *t* to powers in C. This makes the biochemical function BF *type* different from C, while it also allows us to maintain that the biochemical functions are co-temporally materially dependent on the chemical ones. Biochemical functions can then be considered weakly emergent from the chemical powers of the molecule and this provides an answer to the relation problem: the relation between the chemical properties of a biochemical kind and the functions is weak emergence. This also allows the identification of a relation between structural and functional properties, given by the token identity of the instances of the biochemical functions and the chemical properties, while at the same time maintaining a type difference and the related causal efficacy. Moreover, as will be elucidated in the next subsection, this view is also compatible with MR and MD.

In the case of vitamin B12, **Coenz-Blood** has a specific proper subset of the chemical powers of cobalamin, the ones relevant to the regeneration of erythroblasts in hematopoiesis. Those powers are those involved in the relevant co-enzymatic action that the vitamin plays: the token of the powers of **Coenz-Blood** are the same token powers of the cobalamin compound involved in the process, however the causal contribution is distinctive. The function **Coenz-Blood** emerges from the chemical compound in that it has a proper specific subset of causal powers. Specifically, in this specific case, it amounts to those chemical properties that allow for the regeneration of methionine via "the transfer of a methyl group from 5-methyl-THF to homocysteine via methylcobalamin" (Koury and Ponka 2004: 109). This set is not arbitrarily chosen, but it is identifiable thanks to the evolutionary history of the different biological processes in which B12 acts as a co-enzyme [see Figure 1]. The causal contributions are those relevant to the given environment and the given process. The biochemical functions of B12 vitamins can be considered weakly emergent from the chemical dispositional properties of cobalamin compounds at a given time *t*. This makes the causal profile of vitamin B12 distinctive, as recognised in scientific practice and in the functional characterisation of B12. At the same time, this emergence is only weak as it does not presupposes any stronger forms of ontological novelty, as the one of a strong form of emergence of a physically unacceptable variety. The identity of the token powers associated with both the emergent feature and the lower basis allows us to maintain a relation between structural and functional properties. The proper subset view and weak emergence allow us then to answer to the relation problem.

4.3 Multiple Realisability and Multiple Determination

As previously presented, biochemical functions are multiply realisable, and in some biochemical cases, such as in the crystallin protein, the same chemical features can be determined into many biochemical functions. This is often taken as a challenge to the identification of a relation between structure and function. Here, we have presented the proper subset view and weak emergence as an answer to the relation problem. However, more must be said on how this view can be compatible with MD and MR.

MR and MD are "type issues": it is the realised *type* that can be multiple realisable or be one of the determinations of a given lower-level feature. How are they compatible with weak emergence as defined above? Starting with MR, it is the type function **Coenz-Blood** that is multiply realisable by the four vitamers of B12. However, in a given moment, such as during a specific instance of hematopoiesis, a token instance of **Coenz-Blood** will be realised by a specific token instance of the four vitamers of B12. At the time *t*, *only* the token powers of a proper subset of the lower-level entity are identical to the token powers of the emergent feature **Coenz-Blood**. This implies that despite MR at the type level, at *t* the token entity is realised by one lower-level set of features. In the case of MD instead, there is only one token subset of powers that in a given time *t* realises the biochemical functions under discussion. A token biochemical function is emergent in that it has a proper subset of the token powers of chemical features. This makes the proper subset view straightforwardly compatible with multiple realisation and multiple determination, as discussed by Tahko (2020, 2021). Let us consider these them in more detail.

For MR, there may be several distinct token proper subsets of powers of the chemical features that can be associated with the biochemical function. In the case of **Coenz-Blood**, there are several distinct token proper subsets of the B12 vitamers that can be associated with the function and, as such, can realise the biochemical function under consideration. This is possible because, while the type is multiply realised, the token is always realised by a specific subset of token powers. For MD, two aspects should be considered. From the perspective of the token realised feature, one identifiable proper subset of chemical powers is associated with the higher-level feature, and, as such, MD is not problematic. From the multiply determinable feature perspective, instead, the token set of powers of a given chemical feature could present different proper subsets of token powers associated with different biochemical functions. Or, as suggested by Tahko 2020, there could be one proper subset of powers associated with two distinct type features, bringing in different effects in the relevant biological context. Accordingly, the token powers of the functional properties are a subset of those of a single chemical kind [Fig. 2].

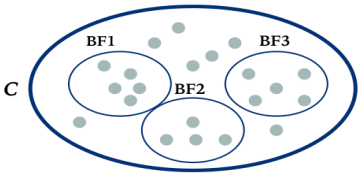


Fig. 2 – Multiple determinability of the cobalamin molecule, for which only one subset of powers is realised at a given *t*.

Moreover, the proper subset view is also able to deal with the reductionist view of MR for which it can be explicated in terms of a closed disjunction. This would make the biochemical functions reducible, and not emergent, to a closed disjunction of chemical structural powers. In this respect, Wilson discusses how the proper subset view ensures a form of ontological autonomy *contra* the disjunctive

strategy (2021). In the case of MR, when the entity H is weakly emergent, the token powers of H are a proper subset of the token powers of either L1 or L2. This makes H type different from the disjunction of Ls because of Leibniz's law: there are some powers of L that are not of H. Moreover, the nature of biochemical functions as defined here also allows to see how the defended view is compatible with MR and MD. The BF is associated with a set of powers whose selection is at least indirectly a result of evolution, and their causal efficacy is embedded in biological systems that are currently evolving. This has an impact on the fact that the types of realisers of the biochemical functions can change or increase in time. In addition to this, there could be a biologically possible world in which the biochemical function is realised by another chemical molecule yet unknown, or that does not play the function in current systems, but could have the function. This would make the disjunction an open disjunction, and, as such, challenges a straightforward reductionist approach.

In conclusion, the proper subset view and an account of weak emergence seem to be compatible with accounting for forms of MR and MD.

5. Conclusion

In this paper, I have considered how biochemical functions can be linked to chemical structure by using Wilson's account of weak emergence (2011, 2015, 2021). Section 2 introduced why the relation between structure and functions in biochemistry is interesting from a philosophical perspective and why can be seen as problematic. Section 3 focused on the double problem of biochemical function, the "function problem" and the "relation problem" offering further context to this debate. Section 4 then explored how, given a definition of biochemical functions, we can interpret the relation between chemical structure and biochemical structure via weak emergence. In doing so, I have considered how this framework offers us a way to think about the relation between structure and function that is compatible with multiple realisability and multiple determinability.

This paper has a series of interesting results. First, it enriches the case studies compatible with Wilson's account of weak emergence. This can bring in new insights relating to the emergence between entities that we would associate to the same level (Bellazzi, 2023). Second, it relates to one of the main research topics of biochemistry, the relation between biochemical functions and chemical structure. The account presented allows us to maintain a form of autonomy for biochemical functions while being compatible with the identification of the relation between structure and function. Third, the results of this paper contributes to the debates on unity of science and reductionism. In particular, they could be further explored to develop the our understanding of the interface between chemistry and biology, if we can establish a relation between the functional and chemical aspects of biochemical kinds.

References

Bartol, J. 2016, "Biochemical Kinds", *British Journal for the Philosophy of Science*, 67, 531-51.

Bellazzi, F. 2022, "Biochemical Functions", *British Journal for the Philosophy of Science* (waiting for an issue).

Bellazzi, F. 2023, *Biochemical Kinds and the Unity of Science*, PhD in Philosophy Thesis, University of Bristol, UK.

Combs, G.F. 2012, *The Vitamins: Fundamental Aspects in Nutrition and Health*, San Diego: Elsevier Science-Technology.

Fang, H., Kang, J., and Dawei, Z. 2017, "Microbial Production of Vitamin B12: A Review and Future Perspectives", *Microbial Cell Factories*, 16, 15, DOI: 10.1186/s 12934-017-0631-y

Gillett, C. 2013, "Constitution, and Multiple Constitution, in the Sciences: Using the Neuron to Construct a Starting Framework", *Minds & Machines*, 23, 309-37.

Goodwin, W. 2011, "Structure, Function and Protein Taxonomy", *Biology and Philosophy*, 26, 533-45.

Havstad, J.C. 2016, "Proteins: Tokens, Types and Taxa", in Kendig, C. (ed.), *Natural Kinds and Classification in Scientific Practice*, New York: Routledge, 74-86.

Havstad, J.C. 2018, "Messy Chemical Kinds", *British Journal for Philosophy of Science*, 69, 719-43.

Koury, M.J. and Ponka P. 2004, "New Insights into Erythropoiesis: The Roles of Folate, Vitamin B12, and Iron", *Annual Review. Nutrition*, 24, 105-31, DOI: 10.1146/annurev.nutr.24.012003.132306

Kistler, M. 2018, "Natural Kinds, Causal Profile and Multiple Constitution", *Metaphysica*, I, 19, 1, DOI: https://doi.org/10.1515/mp-2018-0006

Polger, T.W. and Shapiro, L.A. 2016, *The Multiple Realization Book*, Shapiro, L.A. (ed.), Oxford: Oxford University Press.

Santos, G., Vallejos, G., and Vecchi, D. 2020, "A Relational-Constructionist Account of Protein Macrostructure and Function", *Foundations of Chemistry*, 22, 3, 363-82.

Shapiro, L. 2020, "Theories of Multiple Realisation", *American Philosophical Quarterly*, 57, 1, 17-30.

Slater, M.H. 2009, "Macromolecular Pluralism", *Philosophy of Science*, 76, 851-63.

Tahko, T.E. 2020, "Where Do You Get Your Protein? Or: Biochemical Realization", *British Journal for the Philosophy of Science*, 71, 3, 799-825.

Tahko, T.E. 2021, *Unity of Science*, Cambridge Elements in Philosophy of Science, Cambridge: Cambridge University Press.

Tobin, E. 2010, "Microstructuralism and Macromolecules: The Case of Moonlighting Proteins", *Foundations of Chemistry*, 12, 1, 41-54.

Wilson, J. 2011, "Non-Reductive Realization and the Powers-Based Subset Strategy", *The Monist*, 94, 1, 121-54.

Wilson, J. 2015, "Metaphysical Emergence: Weak and Strong", in Wuthrich, T. (ed.), *Metaphysics in Contemporary Physics: Poznan Studies in the Philosophy of the Sciences and the Humanities*, Schöningh: Brill, 251-306.

Wilson, J. 2021, *Metaphysical Emergence*, Oxford: Oxford University Press.

Emergence, Exclusion, and the Proper Subset of Powers Strategy

Karen Bennett

Rutgers University

Abstract

Wilson characterizes weak and strong emergence partly based on their differing solutions to the exclusion problem. The weak emergentist should claim that emergent phenomena and their bases can both cause the same effect without overdetermining it, because they literally share causal powers. I compare this strategy with a different but related strategy also available to the weak emergentist, and argue that the virtues of the former cost more than it appears.

Keywords: Causal powers, Dependence, Emergence, Exclusion problem, Mental causation, Nonreductive physicalism, Overdetermination.

1. Introduction

Jessica Wilson's *Metaphysical Emergence* (2021) is an excellent and important book that brings together roughly twenty years of work on the ways in which one set of phenomena could be dependent on, and yet to some degree autonomous from, another set of phenomena. Wilson identifies the core shared ideas in the sea of mushy and contradictory usages of the term 'emergence', and articulates notions of 'weak' and 'strong' emergence that (in the philosophy of mind case) correspond to nonreductive physicalism and dualism respectively. She distinguishes these positions, in part, by how they approach the well-known exclusion problem for mental causation. Wilson's discussion of emergence and exclusion will be my focus in this commentary. What exactly does solving the exclusion problem require, and how exactly does her version of weak emergentism pull it off?

Before getting started in earnest, however, I would like to briefly call attention to a particular virtue of Wilson's book: its engagement with, and reliance upon, classic older work in the metaphysics of mind. She engages with a lot of material by people like Terence Horgan, Jaegwon Kim, Andrew Melnyk, Sydney Shoemaker, and Stephen Yablo. This is both appropriate and important, because a lot of excellent work in this area has been somewhat neglected of late. Both Wilson and I began our careers thinking about the mind-body problem, and are therefore well aware that the question of how some things give rise to other things

Argumenta 10, 1 (2024): 237—246
ISSN 2465-2334
DOI 10.14275/2465-2334/202418.ben
First published: 30 November 2024

is not exactly a new topic in metaphysics, as those in the contemporary grounding literature sometimes seem to suggest.

2. Weak and Strong Emergentism, Characterized by How They Handle the Exclusion Problem

Although terms like 'emergence' and 'emergentism' are used in many slightly different ways, Wilson argues that the most basic commitment of philosophical positions worthy of these labels is that emergent properties and states of affairs involve 'autonomy with dependence'. They are synchronically and non-causally dependent on their base, and yet *somehow or other* are autonomous from it: they have different causal powers, figure in different laws, or something along those lines.

That 'somehow or other' is, of course, crucial. Wilson distinguishes two primary forms of emergentism as meaning quite different things by the claim that emergent phenomena have 'different causal powers'. *Weakly* emergent features—if there are any—have fewer causal powers than the bases from which they arise, and *strongly* emergent features—ditto—have more causal powers then their bases. Wilson draws this distinction in the course of exploring available emergentist answers to the exclusion problem. It's a rather neat methodological trick: she simultaneously explains how these two kinds of emergence have different available responses to the exclusion problem, and uses their responses to the exclusion problem to shed light on the difference between them (Chapter 2).

Here's a simple version[1] of the exclusion problem, formulated as a set of five inconsistent claims:

Distinctness: Mental properties (and perhaps events) are distinct from physical properties (events).

Efficacy: mental events cause things, including physical things, and at least sometimes do so in virtue of their mental properties.

Completeness:[2] every physical effect has a sufficient physical cause.

Exclusion: all events that have multiple sufficient causes (that are not themselves causally related)[3] are overdetermined.

Nonoverdetermination: the effects of mental causes are not routinely and systematically overdetermined.

So, the physical effects of mental causes both are and are not systematically overdetermined. No bueno.

[1] The main way in which this version is simplified is that I merely gesture at how it can be run in either or both a property (type) or event version (token). Further, this is not how Wilson presents it. While the differences do not matter to anything of substance, footnotes 4 and 6 are worth reading.

[2] Most people, including Wilson, call this 'closure'. I prefer the label 'completeness', because the term 'closure' suggests that physical effects have *only* physical causes. That is an excessively strong premise that blocks the weak emergentist solution from the start.

[3] The parenthetical clause is there because the proper formulation of Exclusion ought not say that the outcome of a single, non-branching causal chain is overdetermined. If $c_1 \rightarrow c_2 \rightarrow c_3 \rightarrow e$, then e has multiple distinct sufficient causes but is not overdetermined by anyone's lights. An alternate way to circumvent this issue is to instead stipulate that the multiple sufficient causes be direct/unmediated.

One can of course dissolve the exclusion problem by denying that there are any mental phenomena, or claiming that they are epiphenomenal, or insisting that they are to be identified with the physical after all. But, as Wilson points out, these are not *emergentist* responses. They do not respect the core commitments that a) the mental is in some sense emergent (and thus exists), and b) emergent phenomena are in some sense causally autonomous (so mental events/properties are neither epiphenomenal nor identical to their physical bases).

So how should emergentists respond to the exclusion problem? Wilson claims that there are two and only two properly emergentist moves that can be made. The first is to deny Completeness, and claim that mental phenomena have genuinely novel causal powers that are neither determined by nor dependent on their physical bases. This strategy is non-physicalist, and is the distinctively strong emergentist position. The second solution is to deny Exclusion, and say that mental phenomena are causally efficacious and yet their effects are not overdetermined, or at least not overdetermined in the two-kids-simultaneously-throwing-two-rocks-at-a-window variety.[4] This is the weak emergentist or nonreductive physicalist (henceforth 'WE/NP') strategy.

The key WE/NP move is to appeal to an intimate relation short of identity, such as—to borrow Wilson's list (55-57)—functional realization, constitutive mechanism, mereological realization, the determinate-determinable relation, or 'superdupervenience'. (Though Wilson herself would wince (2014, 2018), we might replace some or all of those relations with grounding.)

I have long been fond of the WE/NP response to the exclusion problem, which I once called 'compatibilism'[5] (Bennett 2003). It will be the focus of the rest of the paper.

3. A "Deeper Unity of Strategy"? The Proper Subset Condition and the Counterfactual Condition

Wilson suggests that the fact that different WE/NPs appeal to different intimate non-identity relations is relatively unimportant as far as the exclusion problem is concerned, because

> underlying the seeming diversity in these and many other accounts of nonreductive physicalism hides a deeper unity of strategy (57).

[4] It is just a terminological matter whether we describe this move as saying that the effects of mental causes are not overdetermined at all, or as saying that they are not overdetermined in the bad 'double-rock' way. Discussions and defenses of the strategy take both forms in the literature. (See, e.g., Bennett 2003 and Sider 2003.) Wilson herself frames the strategy in the latter way, as "allowing that [the effects of mental causes] are overdetermined […] but maintain[ing] that the overdetermination here is of an unproblematic *non*-double-rock-throw variety" (44). Characterized like that, the move denies Nonoverdetermination rather than Exclusion: the effects of distinct causes are always overdetermined, but it turns out that overdetermination is more widespread and less troublesome than usually thought.

I prefer the characterization in the main text, which reserves the word 'overdetermination' for the double-rock-style cases. It is also the better characterization for Wilson herself. See note 6.

[5] I called it that because it says that the non-overdeterministic causal efficacy of the mental is compatible with the conjunction of Completeness and Distinctness.

I agree that there is a deeper unity of strategy here. Indeed, I have argued that there *must* be a deeper story, in the sense that the WE/NP ought not simply name an intimate non-identity relation, and announce that events related in that way do not overdetermine their effects. That is not good enough. What is required is a story about how and why that relation has that kind of impact:

> the burden is on the compatibilist here. She needs to be able to *argue* that the effects of mental causes are not overdetermined, and to explain *why* they are not (2003: 474).

That is, in essence, what Wilson is after when she claims a "deeper unity of strategy". She is saying that all of the tight relations postulated by the WE/NP lend themselves to a particular sort of explanation: what I hereby dub the "Proper Subset Strategy".

While I clearly agree about the need for *some* kind of deeper explanation, I am not convinced that the Proper Subset Strategy is the right one. An alternative is available whose relative merits must also be investigated. After sketching both Wilson's story and this alternative, I will explore the relation between them, and argue that the apparent virtues of the Proper Subset Strategy cost more than it seems.

4. Wilson's Proposed Underlying Idea: The Proper Subset Strategy

Wilson claims that whenever one phenomenon *E* is weakly emergent from a base phenomenon *B*, *E*'s causal powers will be a non-empty proper subset of *B*'s. In particular, when mental and physical phenomena stand in any of the close relations posited by the WE/NP, it will be the case that mental phenomena have a non-empty proper subset of the causal powers of the physical phenomena from which they weakly emerge (58-66). Thus the various particular mechanisms for securing weak emergence "are unified in each [endorsing the Proper Subset Strategy] as a means of avoiding problematic overdetermination" (66).

The Proper Subset Strategy certainly sounds good. Indeed, it sounds like it decisively solves the exclusion problem. The picture is that mental and physical causes do not overdetermine their effects because there is a *literal shared core* of causal juice: to say that mental phenomenon *M* and its physical base *P* overdetermine their effects would be wrong in the same way that it would be wrong to say that our two favorite hooligans, Billy and Suzy, overdetermine the breaking of the window by holding hands and jointly throwing one single mutually-owned rock. It is wrong in the same way that it would be to say that you and I double-pay the bridge toll by together tossing in one $5 bill from our shared piggybank, or that it would be to say that there are two winners of the local 5K, the Johnson family and the García family, because Inez García-Johnson won it. In none of these cases is there any genuine doubling. The window's breaking has just one proximate cause; the 5K has just one winner; the bridge toll has been paid only once. Exclusion begone![6]

[6] Now it can be seen that it is less than optimal for Wilson to characterize the WE/NP solution to the exclusion problem as saying that the effects of mental causes *are* overdetermined, but not in the bad double-rock way—that is, as denying Nonoverdetermination

Unfortunately, this is all a bit of legerdemain. But before I explain why, I need to put the alternative on the table.

5. An Alternative Underlying Idea: The Counterfactual Strategy

Talk of overlapping sets of causal powers is not the only way to explain how various intimate relations between the causes defuse the threat of overdetermination. In a 2003 paper, I offered a different explanation. I provided a necessary condition on overdetermination (genuine, 'double-rock' overdetermination), and argued that it is not met by pairs of causes related in any of the ways WE/NPs think that mental and physical phenomena are.[7]

The necessary condition is simply that two causes overdetermine an effect only if had either happened without the other, the effect would still have occurred.[8] That is, causes c_1 and c_2 overdetermine e only if both of the following counterfactuals are nonvacuously true:

$(c_1 \& \sim c_2) \mathbin{\square\!\!\rightarrow} e$

$(c_2 \& \sim c_1) \mathbin{\square\!\!\rightarrow} e$

This is a very intuitive test for overdetermination. We implicitly rely on it whenever we distinguish between overdetermination and joint causation. Indeed, note that those who would appeal to modal fragility to claim that all apparent overdetermination is really joint causation implicitly rely on these counterfactuals.[9]

Yet if the test is legitimate, the WE/NR is again in good shape. At least one of these counterfactuals will be vacuous or false when (2003) and only when (2008) the mental and physical causes stand in one of the WE/NR's favored relations. Though the details get too complicated to revisit here, the basic idea is that on any such relation, the physical base necessitates the weakly emergent mental phenomena, rendering one of the counterfactuals vacuous.

6. The Relation Between the Two Strategies

Two ways of explaining why the existence of certain tight relations falsifies Exclusion are now on the table. Each strategy offers a necessary condition on overdetermination—one, that certain counterfactuals be nonvacuously true; the other, that the two potential causes not be such that one's set of causal powers is a proper subset of the other's—and claims that weakly emergent phenomena and their

rather than as denying Exclusion. (See note 4). Given the Proper Subset of Powers strategy, she should not think that the effects of mental causes are overdetermined *at all*. For an effect to be overdetermined, it must have at least two distinct causes. But the only sense in which Wilson's WE/NP thinks there are two distinct causes is that there are two distinct phenomena that literally share the efficacious part.

[7] Really, in any of the ways *any* physicalist thinks they are: identity works too.

[8] This is not supposed to be an analysis of overdetermination in noncausal terms, just a condition on which causes count as overdeterminers.

[9] Billy and Suzy throw separate rocks, apparently overdetermining the breaking of the window. The fan of the fragility treatment of such cases (Lewis 1986, 2000) would say, "look I know it seems like the window would still have broken if only Billy threw his rock, or only Suzy threw hers. But that's not actually true, because the precise time and manner of the breaking are essential to it. If only one of them had thrown, it would not have been the very same break. So you're wrong about those counterfactuals. The particular window-breaking that actually happened required both Billy and Suzy to throw their rocks".

bases do not meet the condition, and thus do not overdetermine their effects. Here is a bit more about the relation between these two conditions.

First, the failure of the causal powers to nest in a subset relation does not entail that the overdetermination counterfactuals are nonvacuously true. There are at least two reasons for this. One is that someone who denies that there are any such things as causal powers, or that (foreshadowing!) they are the kinds of countable things that can form sets, will deny that any pairs of events are such that their causal powers nest in the relevant way. But such a person is not committed to thinking that all overdetermination counterfactuals, formulated with whatever pair of events you like, are nonvacuously true. Another reason is the case in which c_1 and c_2 share a lot of causal powers, but not all of them; the two sets overlap but neither is a subset of the other. It could still be the case that one or both of the overdetermination counterfactuals is false or vacuous, for example if the non-shared causal powers are irrelevant to the particular effect in question.

What about the other direction? Does the nonvacuous truth of the overdetermination counterfactuals entail that the causal powers fail to nest in a subset relation? Equivalently, does the subset-nesting of the causal powers entail that at least one of the corresponding overdetermination counterfactuals is false or vacuous? It is tempting to say yes, but matters are somewhat tricky.

Suppose that c_1's causal powers are a proper subset of c_2's, and that c_1 and c_2 are both actual causes of e. It is likely nonvacuously true that if c_1 had happened without the 'larger' c_2, the effect would still have happened. The interesting question is whether e would still have happened if c_2 had happened without the 'contained' c_1. The difficulty in assessing the counterfactual is that the mere claim that c_1's causal powers are a proper subset of c_2's says nothing about the modal status of that inclusion, nor about whether either event has any or all of those token causal powers essentially. The whole shebang could be contingent. And that makes it difficult to mount a decisive case for the falseness or vacuity of the overdetermination counterfactual $(c_2 \,\&\, \sim c_1) \,\square\!\!\rightarrow e$. The options are that a) c_2 cannot happen without c_1, in which case the counterfactual is vacuous, b) c_2 *can* happen without c_1, and indeed with c_1 and all its causal powers deleted completely, in which case the same counterfactual is probably false, and c) c_2 can happen without c_1 in particular, but only if c_1's causal powers are replaced by numerically different but qualitatively similar ones (in the way that an object might survive the replacement but not complete loss of a part). In *that* case, the counterfactual is probably nonvacuously true, despite the 'subsetting'. And this is the most likely case in the situation at hand: where c_2 weakly emerges from c_1, via any of the standard WE/NP relations. Maybe this mental state could happen without this *particular* physical state that underwrites it, but it cannot happen without *any* physical basis.

Now, I do not want to rest a lot of weight on this. I myself have argued that these kind of 'replacement' interpretations of counterfactuals are problematic (2003: 482), and David Lewis seems to agree (2000: 190). My only point here is that the path from causal-power-subsethood to the falseness or vacuity of the overdetermination counterfactuals is neither obvious nor straightforward. Given the entailment failure in the other direction, it is probably best to think of the two strategies as independent. Two events that vacuify or falsify the counterfactuals need not meet the Proper Subset Condition, and it may well be that two events that meet the Proper Subset Condition can fail to vacuify or falsify the counterfactuals.

7. The Proper Subset of Strategy Is Not More Powerful than the Counterfactual Strategy

I have sometimes thought that the Proper Subset Strategy is a more powerful (groan) implementation of the Counterfactual Strategy. (Both appeared in print at roughly the same time: e.g. Wilson, 1999, 2002; Shoemaker 2001, 2003; Bennett, 2003.) I have come to think that this is wrong. The previous section shows that it isn't clearly right to think of the Proper Subset Strategy as an implementation of the Counterfactual Strategy. And although there is a clear case to be made for the claim that it is more powerful, in two specific senses, this advantage is an illusion.

The first sense in which the Proper Subset Strategy seems to be more powerful than the Counterfactual Strategy is that it appears to provide a deeper, more convincing explanation of why there is no overdetermination. Recall the examples of the bridge toll, the 5k, and the hand-holding hooligans: the weak emergentist gets to similarly claim literally shared causal power. In contrast, the Counterfactual Strategy just says something kind of wishy-washy about the truth-values of certain counterfactuals, while remaining silent about *why* those counterfactuals have the truth-values they do.

The second way in which the Proper Subset Strategy seems to be more powerful than the Counterfactual Strategy is that it not only shows that the weakly emergent entities and their bases can both be causally efficacious without overdetermining their effects, but also shows that weakly emergent phenomena are causally efficacious in the first place. If such phenomena have a nonempty proper subset of the causal powers of their bases, then *a fortiori* they have causal powers.[10] The Counterfactual Strategy, in contrast, does not do this. It simply *assumes* that the mental is causally efficacious, and shows that this (together with Distinctness and Completeness) does not entail that the effects of mental causes are systematically overdetermined.

Unfortunately, these two seeming advantages are just that: mere seemings. There is little substance to either point, which I will address in reverse order.

First, a solution to the exclusion problem that establishes the causal efficacy of the mental, or the weakly emergent more generally, is actually not superior to one that does not—at least, not *qua* solution to the exclusion problem. The exclusion problem is an attempt to undermine the causal efficacy of the mental (the emergent), not because of any intrinsic defect, but rather because there is no causal work for it to do.[11] An adequate response to the exclusion problem is simply one that undercuts this reasoning. My point here is just the elementary one that objecting to an argument that $\sim p$ does not require showing that p is *true*. Thus the fact that the Proper Subset Strategy secures the causal efficacy of the mental does not add anything *qua response to the exclusion problem*.

[10] Wilson admits that nothing she says gives the weakly emergent phenomena *novel* efficacy (58, 67-69), but she is right to accept this consequence. It's what makes weak emergence different from strong emergence. No nonreductive physicalist, for example, should grant causal powers to the mental that aren't possessed by its physical base.

[11] Contrast, for example, Princess Elisabeth-style complaints about substance dualism, where the problem is that the mental is not spatially located, has no mass, has no chemical structure, and so forth.

Of course, this does not mean that it is no advantage at all to the Proper Subset Strategy. It could solve the exclusion problem *and* secure the causal efficacy of the mental. But I am still skeptical; I do not think the strategy actually does secure that. All the work is done by Wilson's claim that weakly emergent entities have a *non-empty* proper subset of the causal powers of their bases. This is the only reason we are guaranteed that weakly emergent entities have causal powers. But Wilson never argues that any particular thing or kind of thing has a non-empty set of causal powers; that is just part of her definition of weak emergence. So those who are inclined to be worried about the causal efficacy of the kinds of phenomena she takes to be weakly emergent—like the mental—will simply deny that they are weakly emergent in her sense.

Second, I also doubt that the Proper Subset Strategy truly provides a deeper, more convincing explanation of why there is no overdetermination—no "causal competition" as Yablo puts it (1992). It looks like it does, yes, but, well, that is the nature of prestidigitation.

The problem is that the deeper explanation requires being quite literal about something that it is not so easy to take literally. The way the Proper Subset Strategy so cleanly escapes overdetermination is by *identifying* each and every causal power of the weakly emergent phenomenon with a causal power of the base phenomenon. As Wilson has emphasized since she began defending the view (1999, 2002), it is crucial that each individual causal power of the emergent thing be possessed by both.

To bring this out clearly, consider two similar but hopeless positions that result from removing the 'subset' part from the Proper Subset Strategy. One position simply says that weakly emergent phenomena have fewer causal powers than their bases. This is no help with exclusion at all; a rock presumably has fewer causal powers than a similarly sized iPhone—for example, only the latter can call an Uber—but throwing both can certainly overdetermine the breaking of a window. The second hopeless position says not only that weakly emergent phenomena have *fewer* causal powers than their bases, but also that their causal powers are *qualitatively indiscernible* from those of their bases. But this again is no help with the exclusion problem. Events with non-identical but qualitatively indiscernible causal powers can absolutely overdetermine things. Consider a scenario in which Billy and Suzy stand 5 feet from each other and throw two indiscernible rocks in indiscernible ways at the window, hitting *almost* the same spot with the same force, at the same angle, at the same time. Their rock-throwings share almost all their causal powers at the type level. (That is, the vast majority of the causal powers belonging to Billy's throw are qualitatively indiscernible from those belonging to Suzy's throw.) But the causal powers of the two events are not numerically identical, and their breaking the window is, again, an uncontroversial case of overdetermination.[12]

In short, the success of the Proper Subset Strategy entirely depends on the idea that the causal powers of the emergent phenomena are numerically identical to the causal powers of the base. And this in turn requires that token causal powers

[12] At this point, one might move to the idea that the causal powers of the base *constitute* or *realize* the distinct but qualitatively indiscernible causal powers of the weakly emergent phenomena. This is basically Derk Pereboom's view (2002, 2011). Whatever its merits, it does not avail itself of the Wilson-Shoemaker idea that there is a shared core of causal power.

are the sort of thing that can not only be *counted* but also *individuated*. Indeed, it is very, very hard not to imagine them as pebbles in a bucket—and Wilson's diagrams on page 70 suggest that she cannot resist this picture either. But this is a serious and rather discombobulating ontological commitment. I will not argue here that causal powers are not like that, but I suspect others will share my reticence. Even Wilson takes pains to insist that her causal powers are nothing dubious or creepy:

> Talk of powers is simply shorthand for talk of what causal contributions possession of a given feature makes […] to an entity's bringing about an effect, when in certain circumstances […] no controversial theses pertaining to the nature of powers, causation, properties, or laws are here presupposed (32-33; also 45).

But the question is, can she really make good on this neutrality? More precisely, can she assuage my ontological qualms while retaining the nice claim that strictly speaking, there is really only one cause of an effect caused both by a weakly emergent phenomenon and its base? That is the challenge I lay before her.

Let me be crystal clear: I have not argued that she cannot meet this challenge. I have simply *raised* the challenge. My real point here is that one cannot have the Proper Subset Strategy on the cheap; the cost-benefit analysis must be made. We can shoulder the ontological commitment to trackable, countable causal powers and accept the benefits, or we can be squeamish and reject the whole picture. What we cannot do is help ourselves to the lovely solution to the exclusion problem while acting as though it costs no more than simply believing in causation. When I accuse the Proper Subset Strategy of sleight of hand, that is what I really mean: not that it cannot fulfill its promise at all, but rather that it hides the expensive machinery required to do so. Regardless, I have appreciated the opportunity to drill deeper into it than I previously have, and discover its secrets.[13]

References

Bennett, K. 2003, "Why the Exclusion Problem Seems Intractable, and How, Just Maybe, To Tract It", Noûs, 37, 471–497.

Bennett, K. 2008, "Exclusion again". In J. Howhy & J. Kallestrup (eds.). Being Reduced, 280–305, Oxford: Oxford University Press.

Lewis, D. 1986, "Postscripts to 'Causation'". In his Philosophical Papers, Volume II, 172–213.

Lewis, D. 2000, "Causation as Influence", The Journal of Philosophy, 97, 182–197.

Pereboom, D. 2002, "Robust Nonreductive Materialism", The Journal of Philosophy, 99, 499–531.

Pereboom, D. 2011, Consciousness and the Prospects for Physicalism, Oxford: Oxford University Press.

Shoemaker, S. 2001, "Realization and Mental Causation", in Identity, Cause, and Mind: Philosophical Essays, Oxford: Oxford University Press, 427-451.

[13] Thanks to Jessica Wilson and the audience at the 2023 Eastern APA Author-Meets-Critics for discussion. I was tempted to somehow work the phrase "metaphysical emergency" into the paper, but I resisted. You're welcome.

Shoemaker, S. 2003, “Realization, Microrealization, and Coincidence”, Philosophy and Phenomenological Research, 67, 1-23.

Sider, T. 2003, "What’s so Bad About Overdetermination?", Philosophy and Phenomenological Research, 67, 719-726.

Wilson, J. 1999, "How Superduper Does a Physicalist Supervenience Need to Be?", The Philosophical Quarterly, 49, 33-52.

Wilson, J. 2002, "Causal Powers, Forces, and Superdupervenience", Grazer Philosophische Studien, 63, 53-78.

Wilson, J. 2014, "No Work for a Theory of Grounding", Inquiry: An Interdisciplinary Journal of Philosophy, 57, 535-579.

Wilson, J. 2018, "Grounding-Based Formulations of Physicalism", Topoi, 373, 495-512.

Wilson, J. 2021, The Metaphysics of Emergence, Oxford: Oxford University Press.

Yablo, S. 1992, "Mental Causation", The Philosophical Review, 101, 245-280.

A Mereology for Emergence

Claudio Calosi

University of Venice

Abstract

The paper first investigates the tension between reductive accounts of mereological structure and emergence as characterized in Jessica Wilson's seminal work. It then suggests a new mereology for emergence. Finally, the resulting account is applied to a paradigmatic case of an emergent whole.

Keywords: Emergence, Mereological structure, Mereological sum, Matter.

To my partner in crime, J.W.

1. Emergence and Mereological Reductionism

There are several broadly "reductive" accounts of mereological structure. They all try to capture rigorously the somewhat vague intuition that "wholes are nothing over and above their parts". The most radical view in the reductive camp holds that mereological composition is strict numerical identity, in that wholes are numerically identical to their parts considered collectively. The view is known as *Strong Composition as Identity*. Using double signs (such as xx), for plural terms:[1]

Strong Composition as Identity (CAI): If the xx compose y, then $xx = y$.

There is a famous argument in the literature against CAI from the possibility of emergence.[2] It goes roughly as follows. If CAI is true, then wholes cannot have properties that the plurality of their proper parts do not have. Emergent properties are exactly an example of such properties. Hence, if (possibly) there is emergence, CAI is false. Whatever one thinks of the argument, CAI is indeed a radical option. For example it might require substantive changes in the logic of identity and/or comprehension principles of plural logic. Hence, it is important to realize that the tension between reductive accounts of mereological structure and (the possibility of) emergence cuts a little deeper. As Wilson (2021) puts it,

[1] For an introduction see Baxter and Cotnoir 2013.
[2] See e.g., McDaniel 2008, Schaffer 2010, Sider 2013, and Calosi 2016.

Argumenta 10, 1 (2024): 247—255
ISSN 2465-2334

DOI 10.14275/2465-2334/202418.cal
First published: 30 November 2024

> It is the *coupling of cotemporal material dependence* with *ontological* and causal *autonomy* which is most *basically definitive of the notion of emergence*, at least as suggested by the central cases of special-science entities with respect to the physical micro-configurations which are their constant companions (Wilson 2021: 1; italics added).

In the light of this, the general threat coming from emergence to reductive accounts of mereological structure is the following. If emergent wholes are *ontologically autonomous* from their (microscopic) constituents,[3] then they are indeed "something over and above" those constituents, contra the spirit, not just the letter, of reductive accounts. It is not my purpose here to respond to the threat, nor to dissect its presuppositions. Rather, it is to take such a threat at face value and propose a new mereological system that vindicates the claim that "wholes are something over and above their parts"—as seems to be required by metaphysical emergence. This is by no means an easy task. Indeed, many think that mereology alone is not enough to account for complex, highly structured, emergent wholes. This is why they recommend different forms of hylomorphism.[4] Others think that we need to revisit the very mereological framework we use, for example adopting a so-called slot-mereology,[5] or rejecting mereological monism, roughly the view that there is only one notion of (mereological) part.[6] I am going to suggest a mereological account that uses only one notion of parthood. In a nutshell, I am going to suggest that we can define a notion of mereological sum that is not equivalent to extant ones in the literature. Given anti-symmetry of parthood, it turns out that sums are unique. I then define the notion of the matter of an entity as *the* sum of its proper parts. This helps me draw a distinction between *Reducible Wholes*, wholes that are nothing over and above their matter, and *Irreducible Wholes*, wholes that are distinct from their matter. Finally, I suggest that if a whole is an emergent whole, then it is an irreducible whole—as previously defined.[7]

2. A New Mereology

There are three notions of mereological sum in extant literature.[8] I will use $<$ for parthood, $\ll$ for proper parthood, $\circ$ for overlap, defined as usual, and $\prec$ for the plural logic relation of "being one of".[9] For the sake of readability "$xx < y$" abbreviates "$\forall x(x \prec xx \rightarrow x < y)$", and "$x \circ xx$" abbreviates $\exists y(y \prec xx \wedge x \circ y)$". Then the usual notions of sum are defined as follows:

D. 1 $Sum_1(xx, y) \equiv \forall x(x \circ y \leftrightarrow x \circ xx)$ SUM_1

D. 2 $Sum_2(xx, y) \equiv xx < y \wedge \forall x(x < y \rightarrow x \circ xx)$ SUM_2

D.3 $Sum_3(xx, y) \equiv xx < y \wedge \forall x(xx < x \rightarrow y < x)$ SUM_3

[3] I follow Wilson (2021: 10) here. Roughly, an emergent whole is a whole with an emergent feature.
[4] See e.g., Koslicki 2008, Fine 2010, and Sattig 2015.
[5] See e.g., Bennet 2013 and Sattig 2021.
[6] See e.g., Canavotto and Giordani 2020.
[7] I developed the technical work on the new mereological system together with Alessandro Giordani. See Calosi and Giordani 2023a, and Calosi and Giordani 2023b.
[8] See Cotnoir and Varzi 2021.
[9] That is, $x \ll y \equiv x < y \wedge x \neq y$, and $x \circ y \equiv \exists z(z < x \wedge z < y)$.

In plain English, y is a Sum_1 of the xx iff it overlaps all and only the things that the xx overlap, y is Sum_2 if every xx is part of y and every part of y overlaps the xx, and finally, y is a Sum_3 iff every xx is part of y, and everything that includes the xx includes y. It is well-known that in mereologies that are weaker than classical mereology, the three notions are not equivalent.[10] Do they exhaust the notions of Sum definable in terms of $<$ and $\prec$? Hardly so. Consider the following:

D.4 $Sum(xx, y) \equiv xx < y \wedge \forall x(\neg xx \circ x \rightarrow \neg x \circ y)$
$\wedge\ \forall x(xx < x \rightarrow y < x)$ SUM

Definition **D.4** simply says that y is the Sum of the xx iff (i) the xx are part of y, (ii) whatever is disjoint from the xx is disjoint from y, and (iii) everything that includes the xx includes y. In other words, according to (i), the mereological sum of a plurality should be inclusive enough to count every member of xx as a part. According to (ii), it should be no more inclusive than that. Finally, according to (iii), a mereological sum should be minimal, in that it has to be part of everything that includes the original plurality. It is easily seen that, in the absence of strong mereological principles we have (1) and (2) below, where i ranges over the three notions of sum in **D.1-D.3**:

(1) $Sum(xx, , y) \rightarrow Sum_i(xx, y)$
(2) $Sum_i(xx, y) \nrightarrow Sum(xx, , y)$

Thus, Sum is strictly stronger than any Sum_i. Once we have such a stronger notion of Sum, we can put forward an explicit mereological system based on that notion.[11] For the sake of simplicity, I am going to require a very strong principle for the existence of Sum-s. In particular I am going to require a counterpart of the *unrestricted composition* principle of classical mereology.[12] It should be noted however that weaker principles will do as well. I will return to this shortly. Here is the system:

P.1 $x < y \wedge y < x \rightarrow x = y$ ANTISYMMETRY
P.2 $x < y \wedge y < z \rightarrow x < z$ TRANSITIVITY
P.3 $x \ll y \rightarrow \exists w \exists z(w \ll y \wedge z \ll y \wedge \neg w \circ z)$ QUASI-SUPPLEMENTATION
P.4 $x \prec xx \rightarrow \exists y\big(Sum(xx, y)\big)$ UNRESTRICTED SUM

Let us define "being mereologically simple" and being "mereologically composite" as usual:

D.5 $S(x) \equiv \neg \exists y(y \ll x)$ SIMPLE
D.6 $C(x) \equiv \neg S(x)$ COMPOSITE

It is an interesting feature of the system, and one that is crucial for the present argument, that we have extensionality of Sum, in that Sum-s are unique, but we do not have extensionality of proper parthood. That is, (3) below is a theorem but (4) is not:

(3) $Sum(xx, y) \wedge Sum(xx, z) \rightarrow y = z$
(4) $C(x) \vee C(y) \rightarrow \big((z \ll x \leftrightarrow z \ll y) \rightarrow x = y\big)$

It remains to be seen how this relates to emergence. I now turn to that.

[10] See Cotnoir and Varzi 2021.
[11] This is the system we analyze in detail in Calosi and Giordani 2023b.
[12] Note that REFLEXIVITY ($x < x$) follows.

3. The Account

Given UNRESTRICTED SUM and theorem (3) we can define a total function over the domain of concrete objects that assign to each concrete object its *matter*.[13] More precisely, letting xx be the plurality of proper parts of x, we define the matter of x, $m(x)$ as x if x is simple, and as the Sum of the xx if x is composite:

D. 7 $S(x) \rightarrow m(x) = x$ SIMPLE-MATTER

D. 8 $C(x) \rightarrow m(x) = \iota z\big(Sum(xx, z)\big)$ COMPOSITE-MATTER

Now we can distinguish those objects that are identical to their matter and those that are not. I call the first REDUCIBLE WHOLES, the second IRREDUCIBLE WHOLES:[14]

D. 9 $R(x) \equiv x = m(x)$ REDUCIBLE-WHOLE

D. 10 $I(x) \equiv x \neq m(x)$ IRREDUCIBLE-WHOLE

Intuitively, this distinction corresponds to the distinction between objects that are nothing over and above their parts, such as e.g., heaps of sands, and objects that are something over and above their parts, e.g., complex structured objects such as table, trees, organisms, statues. The following are immediate consequences:

(5) $S(x) \rightarrow R(x)$

(6) $I(x) \rightarrow C(x)$

None of the converses hold. As a way of illustration, consider the following model, where $\oplus$ is simply "binary Sum":[15]

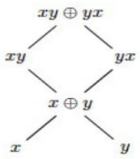


Figure 1: A Model with Reducible and Irreducible Wholes

In the model above $x \oplus y$ is a reducible whole, which is the matter of two irreducible wholes with reducible proper parts, namely xy, and yx, and the matter of a reducible whole with irreducible parts, namely $xy \oplus yx$. It should be clear why the present proposal has a chance to provide a mereology for emergent wholes: it allows for irreducible wholes that are something over and above

[13] As I pointed out before, I require **P.4** only for the sake of simplicity, but it is unnecessarily strong. All the following arguments require is an existence axiom for *Sum*-s that guarantees that the matter of every entity exists. There are different principles that are (i) are compatible with this requirement, and (ii), weaker than **P.4**.

[14] This mirrors the distinction between *unstructured* and *structured* entities in Calosi and Giordani 2023a.

[15] In Calosi and Giordani 2023a we suggest this is how to account for the infamous case of the composition of a syllable in Aristotle's *Met. Z.*

their proper parts, i.e., their matter. Indeed, I suggest that, faced with cases of emergent wholes (E) we should endorse the following conditional:

(7) $E(x) \rightarrow I(x)$

IRREDUCIBILITY as defined above is a necessary condition for emergence. I want to stay neutral as to whether the converse holds. Indeed, I am more hesitant to subscribe to irreducibility being sufficient for emergence. Perhaps there are other "grounds" for irreducibility. Why should one hold that emergent wholes are irreducible in the precise way I defined them? To answer this question, note that we can extract different broad conditions a mereology for emergent wholes needs to meet from the account of emergence in Wilson 2021. Irreducibility in this precise sense helps meeting this requirement. We saw the first (conjunctive) requirement already:

Dependence and Autonomy: Emergent wholes are somewhat dependent on their parts, but at the same time somehow ontologically autonomous from them.[16]

In Wilson's words:

> Summing up: many considerations, drawn from science, perception, language, our practices of individuation, and introspective experience, provide prima-facie support for thinking that many broadly natural entities are co-temporally materially dependent on micro-configurations of fundamental physical entities, yet are also ontologically and causally autonomous with respect to these underlying micro-configurations (Wilson 2021: 6-7).

Compositional Flexibility: The existence of an emergent whole depends on the existence of its parts but does not depend on the existence of any specific plurality of proper parts.[17] In effect, the emergent whole is usually taken to be capable of surviving (some) changes in mereological structure—see e.g., Wilson 2021: 6.

Sortal Properties of Ordinary Objects: Some emergent wholes, in particular *ordinary objects*, fall under "sortal features" that do not apply to any collection of proper parts of said wholes and are responsible for their persistence conditions.[18]

To quote Wilson again:

> Candidate sortal features for ordinary objects of the varieties at issue here would be feature expressing membership in the category at issue, such as 'being a table' or 'being a statute' (Wilson 2021: 197).

[16] Wilson (2021) discusses several suggestions to cash out precisely both the *dependence* and the *autonomy* aspects. I will not enter these details here.

[17] It is an interesting question whether this distinction Wilson draws parallels the one in e.g., Simons 1987 between *generic* and *rigid* dependence. My inclination is that both Simons and Wilson are after the same distinction. But the devil is in the details, and I am not sure Wilson would buy the *analysis* of dependence that Simons (1987) puts forward.

[18] Wilson dedicates the entire Chapter 6 to such objects, arguing that they provide an example of Weak Emergence. For Weak Emergence, see Wilson 2021, especially Chapter 3.

As I pointed out already, I want to make a case for the following claim: the mereological system I proposed helps in satisfying all the desiderata above. Consider *dependence*. According to (7), every emergent whole is an irreducible whole, that is, a whole that is distinct from its matter. But note that the *matter* of an irreducible whole is a very *sui-generis* proper part of that whole. In particular it its *only maximal, unsupplemented proper part*. By this I simply mean that every other proper part of the emergent whole is a proper part of its matter, and therefore overlaps its matter. This captures an important sense in which every irreducible whole *depends* on its matter: were we to annihilate its matter, it is unclear that anything would remain of the whole. Note that it is exactly this kind of considerations that are usually taken to be a litmus test for dependence. At the same time, an irreducible whole is *distinct* from its matter. Now, I grant that numerical distinctness is not sufficient for *autonomy*, but I submit, it is at least *necessary*. What about *compositional flexibility*? There is a raging debate over whether mereological sums can undergo mereological changes. But irreducible wholes are exactly those wholes that are *not* Sum-s. Whatever stance one takes on the possibility of Sums of surviving mereological changes, this does not affect the possibility of irreducible wholes to survive such changes. Indeed, the model in Figure 1 shows that different irreducible wholes, such as xy and yx, can have the same matter. Granted, this does not show that the same irreducible whole can have a different matter at different times. Unfortunately, to provide a detailed account of such possibility, one would need to dive deep into the metaphysics of persistence. I cannot do justice to such a project here. I rest content at pointing out that the very distinction between irreducible and reducible wholes provides a leeway to account for both compositional dependence and compositional flexibility. Finally, *sortal properties*. The thought here is that once the distinction between an irreducible whole and its matter is in place, one can simply claim that the relevant sortal property such as e.g., "being a statue" applies to the irreducible whole but not to its matter. The case of the statue is indeed instructive. Let me contrast here the analysis provided by the account I put forward in the paper with another account, that is more familiar in the mereological literature. My contention is that the new account is a better fit with metaphysical emergence.

As we saw in §1 emergent wholes seem to be "something over and above their parts" in virtue of their ontological autonomy. The familiar way of cashing out this proposal in the mereological literature is to endorse a *non-extensional* mereological system, that is, a mereological system that does not have (4) among its axioms or theorems. The system we are investigating is one example. But there are others. Arguably, the most popular one since at least Simons 1987 is the one that endorses Sum_1 as its notion of sum, has **P.1** and **P.2** as its axioms, and replaces **P.3** and **P4** with the following respectively:[19]

P. 5 $x \ll y \rightarrow \exists z(z < y \wedge \neg x \circ z)$ WEAK SUPPLEMENTATION

P. 6 $\exists x(x \prec xx) \rightarrow \exists z\big(Sum_1(z, xx) \leftrightarrow \varphi(xx)\big)$ RESTRICTED-COMPOSITION

[19] But there are many others. For an introduction see Cotnoir 2013. One needs restricted composition because Weak Supplementation and Transitivity, together with Sum_1, yield (3) as a theorem. See Varzi 2009.

Importantly, in this mereological system Sum_1-s are not unique. That is, (3) is not a theorem of the system. Now, suppose we have a statue, call it $Statue$, that is made out of a lump of clay, call it $Lump$, that has two parts, $Lefty$ and $Righty$. According to the more familiar mereological account $Lefty$ and $Righty$ have two Sum_1-s, namely $Statue$ and $Lump$, as in Figure 2 below:

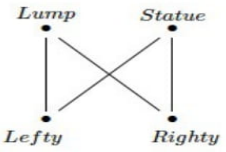


Figure 2: Statue and Lump: Part I

The thought here is that wholes are something over and above their parts in that the existence of proper parts does not determine the identity of the whole. Indeed, different wholes can share the same proper parts. But note that, from a purely mereological perspective, both $Lump$ and $Statue$ are Sum_1 of $Lefty$ and $Righty$. And yet, in the present context, only one of them is an (alleged) emergent whole with a distinguished sortal property such as "being a statue". It seems clear that the mereological structure of the Sum-s cannot account for the difference of the metaphysical status of the wholes with respect to emergence. The mereological system I discussed handles things much differently—and, I contend, better. In the case at hand, there will be only one Sum of $Lefty$ and $Righty$, namely $Lump$ which is a reducible, hence non-emergent whole. $Lump$ is the *matter* of $Statue$ which is a distinct, irreducible emergent whole, as per Figure 3:

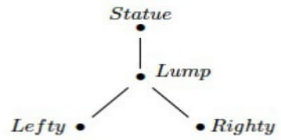


Figure 3: Statue and Lump: Part II

Here, the difference between the composite objects $Lump$ and $Statue$ is reflected in the mereology so to speak. $Lump$ is a Sum, and therefore a reducible object. By contrast $Statue$ is not a Sum. It is something over and above its matter—$Lump$—and this is why the emergent sortal property "being a statue" only applies to *it*. This is reason enough to prefer the mereological system I suggested to the one that is more familiar from the literature, at least if one maintains that statues are emergent wholes distinguished by their emergent (sortal) properties.

4. An Application

Beside ordinary objects and artifacts, Wilson (2021) suggests that special-sciences entities might be (at least weakly) emergent. For instance, she writes:

> Special-science entities are characterized as having distinctive features, constitutive of the distinctive types under which they fall. A tree, for example, has roots, a trunk, branches, stems, leaves; it obtains nutrients from air, sun, soil, and water through leaves and roots; it reproduces via seeds and may bear fruit; it is deciduous or evergreen; it is hardy in certain climate zones, and so on. On the face of it, such features are not appropriately attributed to even complex configurations of fundamental physical entities; and the same is true for the characteristic features of other special-science entities (Wilson 2021: 4).

To conclude I want to discuss an application of the new mereology for emergence that I suggested to a particular example that combines different special-science entities. The example I have in mind is that of the particular *organism* mentioned in the passage above, a tree.[20] How does the new mereology handle the constitution of an organism such as a tree, where different parts of the tree are arguably themselves weakly emergent entities studied by different special sciences?[21] It is interesting to note that the passage to new special-science level with distinctive weakly emergent wholes is clearly mirrored in the mereological system I proposed. In particular it is mirrored in the passage from a reducible whole to an irreducible one of which the former is the matter. For instance, one starts with *atoms*, studied by *physics*.[22] Sums of atoms provide the matter of other weakly emergent wholes, *molecules*, studied by chemistry. Sums of molecules provide the matter for other weakly emergent wholes, *cells*, studied by *biology*. Finally, sums of cells provide the matter of other weakly emergent wholes, *organisms*, studied in the case of a tree, by *botany*. This is illustrated in Figure 4 below:[23]

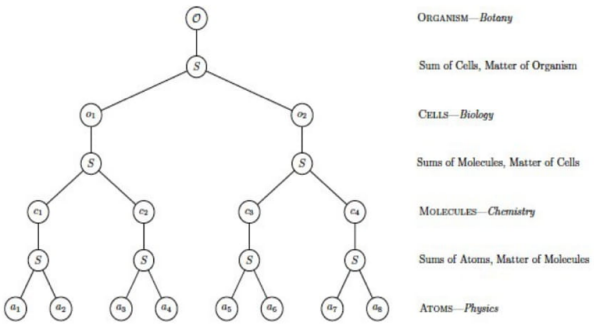


Figure 4: A Tree

[20] See also Calosi and Giordani 2023a.
[21] For a discussion of the relation between emergence, and a layered conception of reality with different levels studied by different special sciences see Wilson 2021: 12 and 24-30.
[22] For a discussion of atomism and emergence see Wilson 2021: 24.
[23] For the sake of clarity, I did not draw all the *Sum*-s.

To sum up. I argued that the possibility of emergence, as characterized in Wilson 2021, poses a threat to various reductive accounts of mereological structure. I then proposed a new account that seems to fit well with various intimations coming from the metaphysics of emergence, as applied to paradigmatic cases of emergent wholes. I admit this is just a first rung of a more thorough investigation of the mereological ladder of such emergent wholes. The hope is that this rung stands on solid ground.[24]

References

Baxter, D. and Cotnoir, A. (eds.) 2013, *Composition as Identity*, Oxford: Oxford University Press.

Bennett, K. 2013, "Having a Part Twice Over", *Australasian Journal of Philosophy*, 91, 1, 83-103.

Calosi, C. 2016, "Composition, Identity, Emergence", *Logic and Logical Philosophy*, 25, 3, 429-43.

Calosi, C. and Giordani, A. 2023a, "Atoms, Combs, Syllables and Organisms", *Philosophical Studies*, 180, 1995-2024.

Calosi, C. and Giordani, A. 2023b, "Universalism and Extensionalism Revisited", https://link.springer.com/article/10.1007/s11229-023-04091-9.

Canavotto, I. and Giordani, A. 2020, "An Extensional Mereology for Structured Entities", *Erkenntnis*, DOI: 10.1007/s10670-020-00305-5.

Cotnoir, A. 2013, "Strange Parts: The Metaphysics of Non-Classical Mereology", *Philosophy Compass*, 8, 9, 834-45.

Cotnoir, A. and Varzi, A. 2021, *Mereology*, Oxford: Oxford University Press.

Fine, K. 2010, "Towards a Theory of Part", *Journal of Philosophy*, 107, 11, 559-89.

Koslicki, K. 2008, *The Structure of Objects*, Oxford: Oxford University Press.

McDaniel, K. 2008, "Against Composition as Identity", *Analysis*, 68, 2, 128-33.

Sattig, T. 2015, *The Double Lives of Objects*, Oxford: Oxford University Press.

Sattig, T. 2019, "Part, Slot, Ground: Foundations for Neo-Aristotelian Mereology", *Synthese*, 198, 2735-49.

Schaffer, J. 2010, "Monism: The Priority of the Whole", *The Philosophical Review*, 119, 1, 31-76.

Sider, T. 2013, "Consequences of Collapse", in Baxter and Cotnoir 2013, 211-21.

Simons, P. 1987, *Parts*, Oxford: Oxford University Press.

Varzi, A. 2009, "Universalism Entails Extensionalism", *Analysis*, 69, 4, 599-604.

Wilson, J. 2021. *Metaphysical Emergence*, Oxford: Oxford University Press.

[24] For comments and suggestions on previous drafts of the paper I want to thank Alessandro Giordani. I also want to thank the editors of *Argumenta*. Needless to say, I owe more than gratitude to Jessica Wilson.

Metaphysical Emergence within Physics: Wilson's Degrees of Freedom Account

Nina Emery

Mt. Holyoke College

Abstract

Metaphysical emergence has often been used to help understand the relationship between the entities of physics and the entities of the special sciences. What are the prospects of using metaphysical emergence within physics, to help understand the relationship between three-dimensional physical entities, and the non-three-dimensional entities that have been recently posited in certain interpretations of quantum mechanics and quantum gravity? This paper explores Jessica Wilson's (2021) analysis of certain cases of metaphysical emergence in terms of degrees of freedom and raises several questions that need to be answered in order to better understand whether this analysis can be used to handle cases of metaphysical emergence within physics.

Keywords: Metaphysical Emergence, Quantum Mechanics, Quantum Gravity.

1. Introduction

In broad strokes, *metaphysically emergent* entities are characterized by being both in some sense *dependent* on some base entities, while also being in some sense *autonomous* from those base entities. Moreover, both the relevant notions of dependence and autonomy are supposed to be suitably *metaphysical*. It isn't enough for the emergent entities to either depend on or be autonomous from the base entities in some merely epistemic or pragmatic sense. Instead, the relevant kind of dependence and autonomy must be understood independently of the kinds of creatures we are, the kind of things we care about, and how we go about investigating the world.

Consider various kinds of special sciences entities—entities that play a role in our best geology and chemistry and biology and so on. On the one hand, the behavior of these entities seems to depend on our best physics; whether you're talking about tectonic plates or chemical solutions or alleles, they are ultimately composed of atoms and subatomic particles (and whatever else physicists turn up in their hunt for a final theory). At the same time, the behavior of entities like tectonic plates and chemical solutions and alleles seems in an important sense

Argumenta 10, 1 (2024): 257—266
ISSN 2465-2334
DOI 10.14275/2465-2334/202418.eme
First published: 30 November 2024

autonomous from the base entities that physics describes. At the very least, we can reliably predict the behavior of these special science entities without paying much attention at all to the details of our best physical theories—indeed that is what geologists and chemists and biologists spend quite a lot of their time doing. Is this type of autonomy suitably metaphysical? It's hard to say, but if it is, then these special science entities would be paradigm examples of metaphysically emergent entities.

So far, so good, but as the reader can surely tell, there's an enormous amount of philosophical work yet to be done both in spelling out precisely what is meant by dependence and autonomy as conditions of metaphysical emergence, and in clarifying when and where in our philosophical theories examples of metaphysical emergence arise. This is the work taken up in Jessica Wilson's important and timely new book, *Metaphysical Emergence* (Wilson 2021). In addition to putting forward a detailed account of metaphysical emergence, Wilson explores the wide range of philosophical arenas in which one might deploy this concept. There are, of course, the standard examples of special science entities mentioned above, as well as the familiar role that emergence has played in the literature on mental causation and causal overdetermination, but *Metaphysical Emergence* also shows how one might use this concept to help think through philosophical questions about the metaphysics of complex systems, ordinary objects, consciousness, and free will.

In this discussion, I'm going to focus on one particular area of application as a way of illustrating both the importance of Wilson's analysis of metaphysical emergence and raising a number of questions about that analysis. In particular, I will be focused on the ways in which the concept of emergence can be deployed *within* physics (as opposed to being deployed as a way of connecting special science entities with the entities of physics, as in the examples above). Wilson discusses this in her chapter on ordinary objects (Chapter 5). But the topic, as I see it, is much more expansive than she has space to take up there.

In recent years, philosophers of physics have gotten quite comfortable with appeals to emergence. Physicists are exceptionally good at generating mathematical formalisms that allow us to make accurate predictions, but the work of interpreting these formalisms—that is, the work of determining what these formalisms tell us about what the world is like—has become increasingly fraught. Often it is the case that the most straightforward or intuitive interpretation of the formalism tells us that the world is dramatically different than we expect it to be—even with respect to the kinds of entities that have traditionally been within the purview of physics. One example of this trend is found in foundations of quantum theory, where some philosophers of physics have begun to advocate for the view that the quantum formalism describes the evolution of a field in an extremely high-dimensional space—a space of 3×10^{80} dimensions.[1] The obvious question that this view raises is how we are supposed to think about the three-dimensional objects that have been the subject of all prior physics—are atoms and the like just an illusion? One way of resolving this question—or at least gesturing in the direction of a possible resolution—is to bring in the concept of metaphysical emergence, and claim that three-dimensional space and the three-dimensional entities occupying that space are metaphysically emergent entities.

[1] See Albert 1996 for an early version of this view and Ney 2021 for a recent comprehensive defense.

A similar line of thought has been highly influential in recent work on approaches to quantum gravity in which there is no spatiotemporal structure.[2] Obviously the world around us appears to have spacetime structure, so doesn't that make these approaches to quantum gravity non-starters? No, the standard line goes, not as long as one is willing to understand spacetime structure as in some sense metaphysically emergent.

These examples show that the concept of metaphysical emergence has the potential to play an important role in philosophy of physics. At the same time, the rules of the game in such debates are very unclear. There is little consensus on the definition or proper analysis of metaphysical emergence among philosophers of physics, or on the more general benefits and challenges of accepting this concept as a part of our overall metaphysical toolbox. Wilson's book therefore should be thought of as providing an important resource to help philosophers of physics think through these issues in a rigorous way that connects with the broader philosophical literature.

2. Wilson's DOF-based Account

As with any account of metaphysical emergence, Wilson's account has two parts: an analysis of the sense in which metaphysically emergent entities are dependent on some base entities, and an analysis of the sense in which metaphysically emergent entities are autonomous from those base entities. The latter is relatively simple (although see more on this in section 5). According to Wilson, the dependence aspect of metaphysical emergence is understood in terms of *cotemporal material dependence*. In paradigm cases (e.g. the special science cases) this involves the base entities *composing* the emergent entity.

The autonomy aspect of metaphysical emergence, on Wilson's view, is understood in two further, distinct ways. In some cases, autonomy is understood in terms of emergent entities having novel *powers* with respect to the base entities. In other cases, it is understood in terms of emergent entities having a proper subset of the powers had by the base entities. Thus we get two types of emergence:

> *Strong Emergence*. What it is for token feature S to be Strongly metaphysically emergent from token feature *P* on a given occasion is for it to be the case, on that occasion, (i) that S cotemporally materially depends on P, and (ii) that S has at least one token power not identical with any token power of P (Wilson 2021: 53).
>
> *Weak Emergence*. What it is for then feature S to be Weakly metaphysically emergent from token feature P on a given occasion is for it to be the case on that occasion, (i) that S cotemporally materially depends on P, and (ii) that S has a non-empty proper subset of the token powers had by P (ibid.: 72).

This classification is all well and good, but I fear that it doesn't help clarify when emergence occurs and when it does not unless we have a settled understanding of powers—when an entity has a power, when it does not, and what precisely powers are. And this, I strongly suspect, is a debate that many philosophers of physics will wish to avoid. With that in mind, it's also important to note that Wilson discusses various "implementations" of weak and strong emergence as defined above, and that one of these—the implementation of weak emergence in

[2] See, for instance, Wüthrich et al. 2021.

terms of degrees of freedom (DOF)—draws on a concept (degrees of freedom) that is already familiar in both physics and philosophy of physics.

Here's how the DOF-based implementation of weak emergence works. As always, the emergent entities need to cotemporally materially depend on the base entites. And then the autonomy condition is understood in the following way:

> […] at least one state of a Weakly emergent entity can be specified using strictly fewer degrees of freedom (independent parameters needed to specify states relevant to an entity's law-governed properties and behaviors) than are needed to specify the corresponding state of the system of entities upon which it cotemporally materially depends (ibid.: 18).[3]

The central example of DOF-based weak emergence, for Wilson is the relationship between the ordinary macrophysical objects that make up the world as we experience it, and the entities described by the quantum formalism. As Wilson writes, "Certain quantum DOF are…eliminated in the classical (macroscopic) limit. For example, entities of the sort treated by classical mechanics are ultimately composed of quantum entities, but the characteristics states of classical-mechanical entities do not functionally depend on the spins of their quantum components" (ibid.: 179).[4]

At least at first, this DOF-based implementation of weak emergence seems highly promising as a tool for understanding emergence within physics. But there are a number of questions that it inspires. In what follows, I'll discuss three of these questions, before returning to briefly discuss Wilson's notion of dependence.

3. The Limits of DOF-based Emergence

Perhaps the most obvious type of question that the introduction of the DOF-based implementation inspires, are questions about the limits of this way of understanding of emergence. First and foremost, we might wonder about the relationship between the DOF-based implementation and Weak Emergence as originally stated. Wilson's presentation of the concept suggests that DOF-based weak emergence only applies in particular cases, where as Weak Emergence is a more general concept. But why, exactly? What are the limits of DOF-based weak emergence? If we wanted to *exclusively* understand weak emergence in terms of the elimination of degrees of freedom, could we? If not, why not?

One way to try to figure out the answers to these questions is by looking at cases where Wilson posits weak emergence without any explicit discussion of degrees of freedom. One especially illuminating example is her application of weak emergence to free will. She writes,

> The prospects [for there actually being free will of the weakly emergent variety] are good. Though free choices are not taken to be part of a higher-level system of laws

[3] Note that Wilson says that the above description is rough. She gives a more thorough, technical definition in chapter 5.2.4. As far as I can tell, however, the details of the technical definition do not affect the discussion here.

[4] Note that although the discussion of ordinary objects being weakly emergent with respect to fundamental particles is the focus of just one subsection of the book (6.1.1), this example is repeatedly mentioned when DOF-based weak emergence is discussed. See, e.g., sections 3.2.3 and 5.2.4.

> on either compatibility or libertarian accounts, a compatibility account is one manifesting the usual Weak emergentist characterization of special science goings on as comparatively insensitive to lower-level physical details, in the sense that an agent's reasons for action in a given case float free of many such details (and in particular, are sensitive only to facts about 'relevant' causal antecedents) (ibid.: 274).

There's no explicit discussion of degrees of freedom here. Why not? One guess is that the mention of laws in the quote above is important. Perhaps on Wilson's view the DOF-based implementation is only possible when the emergent behavior is law-governed. Further support for this guess can be found in Wilson's definition of degrees of freedom. See the quote in section 1 from page 18 and also the following:

> Call states upon which the law-governed properties and behavior of an entity E (object, system, or other particular) functionally depends on the 'characteristic states' of E. A DOF is then, roughly, a parameter in a minimal set needed to describe an entity as being in a characteristic state (ibid.: 177).

From these quotes it looks as though it follows from Wilson's definition of degrees of freedom that if a certain kind of behavior isn't law governed then it won't have any associated degrees of freedom.

This restriction explains the thought that DOF-based weak emergence will only encompass a subset of the cases of weak emergence, but it is a somewhat surprising restriction to make. A fairly standard definition of degrees of freedom is that they are simply the number of independent parameters needed in order to specify a system's state. Of course we tend to only be interested in certain states of certain systems—and therefore we tend to only be interested in certain degrees of freedom. One such group is the states of systems that factor into the laws governing those systems behavior. But there are other salient groups—for instance the states of systems that factor into the explanation of those systems behavior, even if those explanations don't involve laws. And if we have this more expansive understanding of degrees of freedom—where degrees of freedom can be described for any behavior that has an explanation, even if it isn't law-governed—then we should be able to understand compatibilist-style free will as explicitly involving the elimination of degrees of freedom.

All of this by way of discussing how DOF-based weak emergence is related to weak emergence more generally. Another important question about the limits of the DOF-based implementation is whether it can be extended to help us understand strong emergence as well. In the book, Wilson presents this implementation exclusively as a variety of weak emergence. But it seems as though there ought to be a straightforward DOF-based implementation of Strong Emergence, along the following lines:

> *DOF-based Strong Emergence.* There is (i) cotemporal material dependence of the emergent entity on the base entity and (ii) least one state of the emergent entity must be specified using strictly more degrees of freedom than are needed to specify the corresponding state of the system of entities upon which it cotemporally materially depends.

Moreover, at least at first glance, there are some relatively straightforward examples of DOF-based strong emergence in philosophy of physics. For instance, on at least some interpretations of the quantum formalism, when two (or more)

particles become entangled one needs strictly speaking more degrees of freedom in order to specify the behavior of the system than one needs when specifying the behavior of the individual components of the system. For instance, if there are two particles whose spin states are entangled, it may be that all we can say about the behavior of the particles individually is that particle 1 has a .5 chance of having spin up in the z direction and a .5 chance of having spin down, and particle 2 has a .5 chance of having spin up in the z direction and a .5 chance of having spin down. But when it comes to the behavior of the system as a whole, there is an additional important pattern that comes to light, which is that when particle 1 has spin up, particle 2 has spin down. We capture this fact by saying that the wavefunction of the system as a whole takes a certain form, from which it can be derived (using Born's rule) that the probability of the particles having the same spin is 0. A natural way of thinking about this situation is that the entanglement of the particles' spin states results in there being emergent entity—the quantum system—whose state must be specified using strictly more degrees of freedom than are needed to specify the states of the individual particles.

4. Ordinary Objects as an Example of DOF-based Weak Emergence

Another way to try to better understand DOF-based weak emergence is to train a closer eye on some of the examples that Wilson provides. The central example, as mentioned above is ordinary, microphysical objects, which Wilson argues are weakly emergent (in the DOF-sense) from quantum parameters. Here's a bit more of what Wilson says about ordinary objects being weakly emergent.

> What I will call 'classical' objects are ordinary objects of the sort whose static and dynamic behaviors are appropriately treated by classical or Newtonian mechanics, understood as comprising, roughly, Newton's three laws of motion and the gravitational and electromagnetic force laws (ibid.: 192).
>
> The characteristic states of classical objects do not functionally depend on the spins of the quantum components of these entities. Hence notwithstanding that the values of quantum parameters may in some cases lead to macroscopic differences—for example, readings on a measurement apparatus, and the like, as in the case of Schrodinger's cat—it remains the case that DOF such as quantum spin are eliminated…from those needed to characterize entities of the sort appropriately treated by classical mechanics (ibid.: 194).

It is supposed to follow from all this that ordinary objects satisfy the DOF-based account of weak emergence.

The first thing to note about this example is that the details may be dependent on the interpretation that we give of the quantum formalism in fairly complicated ways. Just as one example, in Bohmian mechanics, you can talk about the spin properties of a particle, and use such talk to make predictions, but when you look more carefully, all of the behavior of a quantum particle is explained by its initial position, its initial wavefunction (in the position basis), and the two dynamical laws (the guidance equation and Schrödinger's equation). So it's not entirely clear how to think about the elimination of spin states as a degree of freedom on that interpretation. Was it ever really a degree of freedom to begin with? At the very least there seems to be room for some interesting additional work to be done in

sorting through how this example incorporates the details of various dynamical and ontological interpretations of the quantum formalism.

It's also interesting to note that it isn't immediately obvious why we need to discuss quantum parameters here at all. Consider the fact that ordinary objects like my coffee mug do not unexpectedly lift into the air and float around the room. This behavior is both predicted and explained by classical mechanics. One way of predicting and explaining it is by applying Newton's laws directly to the coffee mug. Another way is to use thermodynamics to predict and explain the behavior of the system involving the coffee cup, the table it is sitting on, and the air around it. Either way, note that you do not need to specify the position and momentum of each individual particle that is a part of the system.

It looks to me like this means that the coffee mug is a weakly emergent entity (on a DOF-based account). The mug cotemporally materially depends on the particles that compose it, but the state of the mug can be specified using strictly speaking fewer degrees of freedom than are needed to specify the states of the individual particles that compose the mug.

Call the argument just given the *classical argument for ordinary objects being weakly emergent* and Wilson's argument described above would be a *quantum argument for ordinary objects being weakly emergent*. At least at first glance it seems that the classical argument works just as well as the quantum argument for Wilson's purposes. And perhaps that's all to the good, since it means we don't have to sort through various interpretations of the quantum formalism in order to conclude that ordinary objects are in fact weakly emergent.

Of course, one thing that seems important about the classical argument is that our best physics says that classical particles with precise positions and momenta are not fundamental. But note first that it wasn't stated in the definition of DOF-based weak emergence that the base entities needed to be themselves fundamental. And second, as mentioned above, it is also controversial whether the quantum entities that instantiate properties like spin and which compose classical objects are themselves fundamental--those who think that the quantum formalism represents a field in a high-dimensional space, for instance, will disagree. So I don't think the non-fundamentality of classical particles is a good reason for treating the classical argument differently from the quantum argument unless you're willing to take a controversial stand with respect to quantum ontology.

4. When Is a Degree of Freedom Eliminated?

It's worth emphasizing the following complication in both the quantum and the classical arguments for the weak emergence of ordinary objects. In terms of the laws governing the base entities, it is *possible* for my coffee cup to lift up off the table and float around the room (or for it to, e.g. quantum tunnel through the table)—it's just very unlikely.

This is importantly different from the example that Wilson gives when discussing what it means for a degree of freedom to be eliminated. In Chapter 5, she writes:

> A case in point is that of a spherical conductor of the sort treated in electrostatics, which has DOF that are eliminated relative to the system of its composing entities; for while the E-field due to the free particles depends on all charged particles, the

> E-field due to a spherical conductor depends on the charges of particles on its surface. Certain quantum DOF are also eliminated in the classical (macroscopic) limit (ibid.: 179).

The case of the spherical conductor is one where degrees of freedom that are in other circumstances relevant to the behavior of the composing entities make *no difference at all* to the behavior of the electric field created by the conductor.

In the classical argument, the degrees of freedom that are in other circumstances relevant to the base entities (i.e. the exact position and momentum of each particle) are *very likely not to affect* the movement of the coffee mug. But there is some probability of them making quite a significant difference. The sense in which quantum degrees of freedom are eliminated in the coffee mug's behavior will also be merely probabilistic. (The exact details of the way in which they are probabilistic will depend on the interpretation one gives of the quantum formalism, but I will try to avoid going too far into the weeds here.)

So one of the key questions facing the DOF-based account is whether that is all that is necessary in order to say that a degree of freedom is eliminated—that it is *very likely not to* have an effect on the behavior of the emergent entity? Another way to put the same point: if a parameter is very likely not to have an effect on the behavior of some entity, is that sufficient to say that the behavior of that entity is *functionally independent* of that parameter?

In part this is just an interesting question to ask about this account. But it also gives rise to an interesting observation, namely that weak emergence might come in degrees, depending on the probability of the "eliminated" degree of freedom actually having an impact on the behavior of the emergent entity. For instance, in both the classical and the quantum case, the probability of a micro-parameter affecting the behavior of an ordinary object will typically decrease as the size of the ordinary object increases. So a larger ordinary object, like a school bus, might be thought of as weakly emergent *to a greater degree* than a smaller ordinary object, like a coffee mug, since the probability of a micro-parameter (e.g. the exact position and momenta of the individual particles) is less likely to affect the behavior of the school bus than the behavior of the coffee mug.

6. What Is Cotemporal Material Dependence?

All of the above discussion has focused on Wilson's understanding of autonomy. Let's turn now to think a bit more about her understanding of dependence. According to Wilson, the type of dependence involved in metaphysical emergence is *cotemporal material dependence*. As noted above, the central examples of emergence (e.g. the special science cases) are cases in which the base entities compose the emergent entities. One would be forgiven, then for thinking that cotemporal material dependence just is composition.

This is relatively straightforward, but it does raise some concerns, in particular about whether and to what extent Wilson's account of emergence can extend to contemporary debates in physics, where it isn't straightforward to understand the base entities as composing the emergent entities. Insofar as one thing helps compose another thing, both entities are standardly assumed to occupy the same physical space. But that assumption breaks down in the examples from philosophy of physics that I introduced at the beginning. If the based entity is a field in a high-dimensional space how can that field composed entities in

ordinary 3-dimensional space? And in interpretations of quantum gravity on which spacetime itself is the emergent entity, it similarly isn't obvious in what sense the base entities would compose the emergent entities.

Comments in the conclusion of the book show that Wilson is aware of this, and is leaving it to future work. That's fair enough, but it's worth pushing a little here, if only to try to get a sense of how this future work is likely to develop.

For instance, in some places in the book, Wilson says that cotemporal material dependence can be "understood as involving both (physical) substance monism and the minimal nomological supervenience of emergent feature types on base feature types" (ibid.: 73). One might take this as an indication that maybe physical substance monism in combination with minimal nomological supervenience is a sufficient condition for cotemporal material dependence.

This is likely to help with the extension of the account to at least some of the contemporary cases in physics. But it does raise some other questions. In particular, it seems like in some cases, composition as an indicator of cotemporal material dependence and minimal nomological supervenience as an indicator of cotemporal material dependence might be in tension. For instance, consider again the cases of quantum entanglement that I suggested in section 2 were potential cases of DOF-based strong emergence. Are these actually cases in which the emergent entity (the entangled system) in fact cotemporally materially depends on the base entities (the individual particles)? It isn't entirely clear.

On the one hand, the entangled system is plausibly composed by the individual particles. But also, the behavior of the entangled system does not nomologically supervene on the behavior of the individual particles—indeed it is the other way around. That's why the case seems like one that would give rise to DOF-based strong emergence.

In fact, if (substance monism plus) minimal nomological supervenience is a sufficient condition for cotemporal material dependence, then maybe cases of entanglement are better understood as cases where the individual particles are *weakly* emergent from the entangled system. After all, on this understanding, the individual particles cotemporally material dependent on the entangled system and you need *fewer* degrees of freedom in order to describe the behavior of those particles.

At any rate, all of this suggests that in order to understand the implications of Wilson's account—and in particular the DOF-based implementation of the account—in philosophy of physics, one will need to not only delve into the complexities of degrees of freedom as indicators of autonomy, but also into cotemporal material dependence as well.

7. Conclusion

The above discussion shows just how rich Wilson's account of metaphysical emergence is by exploring the ways in which just one implementation of her account (the degrees of freedom-based implementation) can be applied to debates within philosophy of physics. The questions raised above are, I think, quite difficult ones. But that just shows how interesting the concept of metaphysical

emergence is and the great potential for important further work on this topic within the philosophy of physics.[5]

References

Albert, D. 1996 "Elementary Quantum Metaphysics", in Cushing, J.T., Fine, A., and Goldstein, S. (eds.), *Bohmian Mechanics and Quantum Theory: An Appraisal*, Kluwer, 277-84.

Ney, A. 2021, *The World in the Wave Function: A Metaphysics for Quantum Physics*, New York: Oxford University Press.

Wilson, J. 2021, *Metaphysical Emergence*, Oxford: Oxford University Press.

Wüthrich, C., Le Bihan, B., and Huggett, N. (eds.) 2021, "Philosophy Beyond Spacetime: Implications from Quantum Gravity", Oxford: Oxford University Press.

[5] Thanks to Jessica Wilson for helpful discussion of the comments at the Eastern APA Author Meets Critics session on *Metaphysical Emergence*, and to Karen Bennett and Brian McLaughlin, who also participated in the session.

The Emerging Limits of Emergentism: Systematicity

Simone Gozzano

University of L'Aquila

Abstract

Taking steps from Wilson's distinction between strong and weak emergence, in this paper I cast doubts on the prospect of weak emergence. After discussing the relationship between properties set at different levels and supporting different counterfactuals and laws, I discuss one crucial condition for a property to be weakly emergent, one that is usually taken as the primary motivation for emergence, that of being "realization indifferent". I set an argument aimed at showing that this realization indifference does not accord with systematic relations holding between properties set at the mental level *vis-a-vis* their realizers. Since it is not possible to have mental properties which are not systematic, mental properties cannot be weakly emergent properties.

Keywords: Emergence, Systematicity, Multiple realization, Realization indifference, Subset.

1. The Making of Emergence

The issue of emergence still is the issue of whether special sciences are autonomous with respect to non-special, or fundamental, sciences. Such an issue was set by the debate, spanned over the years, between Jerry Fodor (1974, 1997) and Jaegwon Kim (1992, 1998 and 1999). The issue of emergence has both an epistemological side—the knowledge and methodology that we use to understand some properties in the world is absolutely specific to those properties?—and an ontological side—are there independent chunks of reality? How do they connect with other chunks?

Thus construed, emergence is seen as an articulated and robust phenomenon. *Articulated* inasmuch there are relations among properties (often called higher-level properties) which are taken to be *independent*, so distinct, of other properties (often called lower-level properties); *robust* inasmuch those relations support counterfactuals, thus allowing for predictions and explanations, that is, for a complex interrelation of epistemological procedures, tenets, and constraints. Or at least those who defend emergence seem to think.

Argumenta 10, 1 (2024): 267—277
ISSN 2465-2334

DOI 10.14275/2465-2334/202418.goz
First published: 30 November 2024

In her book *Metaphysical Emergence*, Jessica Wilson (2021, but see also 2015) argues that we have metaphysical emergence when macro-entities like humans, trees, rocks, and artifacts—as chairs and skyscrapers—are cotemporally materially dependent on but ontologically and causally autonomous from micro-entities, such as quarks and leptons, that ultimately form their base. On this general picture, two varieties of emergence are discussed: weak emergence, which occurs when a high-level feature (be it a property, state or behavior)[1] is both ontologically and causally autonomous and cotemporally materially dependent on a lower-level property or feature—where autonomy is guaranteed by having a subset of the powers had by its base features; and strong emergence, in which along with cotemporally material dependence there is a degree of autonomy to be found in the presence of a new causal power, not to be found in the base features. As such, strong emergence abandons the principle of the causal closure of the physical world, so a high-level feature occurrence cannot be traced back to the occurrence of lower-level physical features. The strong version of emergence proves to be very difficult to defend, while the weak version seems reasonable. But is this the case?

Emergence can be tackled via conceptual analysis and via metaphysics. On the conceptual side, Nicholas Humphreys (2016) has argued along two paths: one is positing that the presence of some properties cannot be derived from the presence of other properties. The other path says that taking certain configurations or patterns as evidence of emergence depends on our conceptualization of those configurations. The first path is conceptual because the notion of *derivation* is not the direct result of the adoption of the nomological-deductive method of science. So, it is a specially tailored notion. The second path, one that applies to phenomena such as flocks of birds or traffic jams, depends on our, presumably *Gestaltic*, capacity of recognizing groups of individual entities moving in a coordinated way as singular entities.

On the metaphysical side, Kim has argued that the nomological relations connecting higher-level *properties*, such as the movement of a flock, could be substituted by lower-level properties, the movements of each bird, thus favoring local reductions. Such local reductions have the burden to show that nothing is lost when the higher-level properties are split into lower-level properties, thus dissolving or reducing the seeming higher-level properties.

How did the attack on special sciences properties develop? One of the attacking points is to consider the predicates used by the special sciences to establish their own domains. For, any new science is characterized by a specific vocabulary, with its predicates and relations. Now, the predicates admissible in laws must be projectible and such that the laws mentioning them support counterfactuals. Being projectible means that the future applications of a predicate are warranted and supported by its past successes. Basically, it is a measure of inductive success, a measure of the force or strength of predicates.[2] Being counterfactual supporting means that the predicates that make a counterfactual true are those that can be included in science because they guarantee the truth of the covering

[1] Somehow betraying Wilson's wording, I will use "features" and "properties" interchangeably.

[2] As a side note in the philosophy of science, one may take it as a sign of resistance to change in science. E.g. "climate change" has not a deep entrenching in past scientific discussions, hence its projectability is modest. Consequently, it is very difficult to take it as a serious player in discussions on the future of climate.

law. Now, the wideness in the support of counterfactuals by a law is a measure of the scope of the application of the law itself. Such wideness can be evaluated both by the number and by the differences of these counterfactuals.

The number of counterfactuals is evidence of how much the law is applied, say, in the same field, thus providing more and more robustness to the projectability of its predicates. The difference in counterfactuals is to be considered in terms of type rather than of token. That is to say, a type different counterfactual establishes specific new relations and it is applied to type different entities and conditions. Clearly, there are cases where there can be type different counterfactuals, and a very high number of them, without this fact providing much insight, as when we say, e.g., that water freezes at 0°C or below and then we may formulate a counterfactual for each fraction of degree below 0°C, which is not very informative. But there are cases in which this number is of interest, as when we consider the angle at which an object bounces in a billiard table or a re-entry trajectory in the atmosphere is to be calculated. Also in this case, we may provide a counterfactual for each value, but the result could prove to be of great importance.[3]

Of greater importance is the number of *type* different counterfactuals supported by a law. Such a number depends on the adaptability of the predicates to new conditions, so by the inductive strength they have. Such strength is made evident exactly by the type-difference of the counterfactuals that the law supports, that is, as said, by the scope of the applicability of the law.

So, the number and types of counterfactuals that a law supports are determined by the strength of the projectability of its predicates, and how much a predicate is projectible depends on the inductive support given by it to successful applications of the law, success measured by the number of conditions in which the law holds. This may sound circular, but since the data and conditions are continuously changing, the circle is not vicious but rather virtuous. In a way, projectability and counterfactual support show us that conceptual analysis and metaphysics are the two sides of the same coin.

It seems, then, that what matters for the inclusion of a predicate into a law is what I would call predicate's *robustness*, namely its projectability and the counterfactual support of the law in which such predicate is included.

One of the most striking examples of this complexity is the way predicates used in psychology are now used in neurology and Artificial Intelligence. Let me contrast three different uses of "is perceiving". 1) A person is visually perceiving satisfactorily if she orientates and navigates herself properly into the world, namely if she finds her way, and does not bump into obstacles. 2) A person is visually perceiving if her eyes, lateral geniculate nuclei, and occipital areas V1-V5 are working and responding to the impinging stimuli determining the appropriate responses from the motor cortex. 3) A robot is perceiving if its cameras, processors, and CPU are such to activate its motor control engine to minimize the number of damaging interactions with the physical world while navigating it appropriately. So, the predicate "x is perceiving the environment" is used in several and type-different ways.[4]

This variety of applications, and this robustness, may come at a cost. On the one side, the wider the application, the wider the projectability and the support

[3] Thanks to Larry Shapiro for having pointed out this problem to me.

[4] I am not getting into the consciousness domain on purpose now, because I do not want to mix the issues.

for a variety of counterfactuals. On the other side, the counterexamples to the inductive base of such large-spectrum predicates can be quite different and revealing of their distinctness. This point was noted by Kim and discussed by Fodor, and the discussion was in terms of potentially disjunctive sets of confirmation.

In their original example, Fodor and Kim were considering "jade": a noun used to refer to two chemically different gemstones, jadeite and nephrite. Now, the sentence "jade is hard" is true both of jadeite and nephrite but this could be the case for different physical-chemical structures.

Fodor stressed that a high-level property could have an open or a closed set of realizers, where it being open is a crucial feature of special sciences. Now, I take the idea of an open set as quite idealized: a set should be closed for it to be defined, so let's say that what Fodor had in mind was an ideally very heterogeneous set. Let's consider pain: supposedly, in humans, it is realized by C-fiber firing, but it could be differently realized in other sentient beings and the realizers form an open set. So, we may take the property of being in pain as one that at a very high level can be shared by different entities, from human beings to other mammals, to other animals up to potentially extra-terrestrial individuals. At a finer level of detail, being in pain is multiply realized by structures that may have nothing in common.

So, is the latency, the wideness in the applicability of predicates and laws, tightly linked and supported by the projectability and number of counterfactuals or should we accept a loose relationship between the underlying (lower-level) structures supporting the higher-level phenomena?

2. Setting a Discussion

The above question bears directly to the issue of emergence, for emergence necessarily entails some form of autonomy between properties (and predicates) as referring to different levels of reality (whatever these levels are). In what follows I will consider weak emergence only, as the strong version seems to have little to no-prospects to be right. Indeed, strong emergentism entails abandoning the principle of causal closure which physicalists take to be non-negotiable. Vice versa, weak emergentism accepts the principle and tries to show that high-level and low-level features do not determine the pernicious overdetermination of so-called double-throw rock variety. Wilson's take on weak emergence is crucially set on the proper subset of power condition (PSPC) according to which a weak emergent feature S has on a given occasion powers that are a proper subset of the powers had by the Ps features on which, in that occasion, S cotemporally materially depend (CMD) (cf. Wilson 2021: 59).[5] In the terms of pain, we may say that John being in pain has both a special science feature (the phenomenal experience John is having) and a physical feature (his C-fibers firing) so that the S CMD on the Ps while being ontologically and causally autonomous from Ps. This PSPC is the way in which this autonomy is spelled out, and such condition is, in a way or another, endorsed or satisfied by all the weak emergentist parties, Wilson argues. This satisfaction, though, comes in different varieties. All these varieties are form of realization. These could be functional, constitutive-mechanistic, mereological, determinable-based or ontologically explanatory realization. Now, some of these varieties of realizations entail multiple realizability: surely functional realization does, but so mereological and determinable-based as well. To wit: one can multiply realize a wall out of the same bricks by

[5] From now on, references to Wilson's book will be just numbers in brackets.

having these parts rearranged (unless endorsing the very much debated constitution as identity thesis) or one may realize red by having either, say, crimson or scarlet and this goes hand in hand with the determinable type having fewer powers than its determinate types, thus satisfying the PSPC (65). Even if I prefer to leave it open whether all forms of realizations entail multiple realizability, we may stress that in most of the central cases of emergence, the way in which the weakly emerging property occurs is indifferent with respect to how it is realized, thus entailing some form of multiple realizability. I will say more on this later on, while defending the second premise of an argument that, I believe, could represent a problem for weak emergence. The argument goes as follows:

(i) Mental features are systematic;
(ii) (Many cases of) Emergence entails realization indifference;
(iii) Systematicity entails that realization indifference cannot hold;
(iv) Therefore, (in many cases) mental features can't be emergent.

3. Defending Premise (i)

We need to defend these premises. One crucial issue is whether mental properties Ss are systematic, as I will argue. That mental properties are systematic can be established via a sort of slippery slope: if some properties are in systematic relations, then you have a lot of systematicity.

Why accept systematicity? For the mental such acceptance is crucial: the more systematic the mental relations are, the less viable a complete reduction of them is. This was Davidson's point (1970) in stressing the anomaly of the mental (and hence its normative nature), or Fodor's (1975) point in stressing the holistic (*Quinean*: each belief is somehow confirmed by every other belief) and deeply inferential (*isotropic*: every belief is somehow pertinent to every other belief) nature of central systems.

The idea of such systematicity is that one can go from one mental state to another via logical or deductive relations. Now, this is surely true of intentional states: assuming rationality (Dennett 1971) or the principle of charity (Davidson 1974) amounts to assuming that crediting one subject with the belief that *p* entails also crediting the subject with those beliefs that follow from *p* at least directly and straightforwardly. Clearly, one has to refrain from assuming logical omniscience, but this can be limited, as I said, by taking only direct inferential links as acceptable. But is that true of qualitative or phenomenal states as well?

I think there are systematic relations also in the case of phenomenal states. Compare two phenomenal state tokens or properties Ss, say the property of feeling pain. We can consider many systematic relations. Let me make two cases for phenomenal states and one in which phenomenal and intentional states are mixed.

From stimulus to phenomenal state: if a subject is stimulated by stimulus R and enters into a phenomenal state S, it could be proved that if the subject receives stimulus 2R (double intensity) it will get into a state nS related to state S by some ratio (as per Weber-Fechner law). So, if these Ss are happening to the same subject along a short interval, we should imagine them being in a mathematical relation that somehow mirrors the values of the stimuli. This relation was supposed to be logarithmic, even if Johnson et al. (2002) have now demonstrated that the basic law of psychophysics vindicates linearity between a subjective experience (or magnitude, as they call it) and the neural activity on which it is based. According to them:

> [the] subjective magnitude, m, depends on a single, unidimensional measure, c, of the complex, multivariate neural response studied in the neurophysiological experiments: m = m(c). [Where] c = c(N), in which c(N) is the function (the operation) that yields the neural coding measure, c. If, for example, c is the mean firing rate of a population of neurons, then c(N) is the operation, summation, required to obtain c (Johnson et al. 2002: 113).

So, a set of phenomenal states, triggered by the same kind of stimuli, present internal relations that can be discovered empirically.

Let me now consider systematic relations among phenomenal states: if the subject gets a phenomenal state S such as to determine some sort of reaction (withdrawal, anxiety) it is natural to imagine that 2S will determine a modification in the speed or intensity of the reactions, even if the amount for such modifications can be hard to determine and may take a lot of empirical work, as happened in the case of the Weber-Fechner law. Again, we can imagine, and we can introspect ourselves to reveal the presence of internal relations between our phenomenal states. If both these cases were to hold, this would be in support of systematicity not only in the case of intentional features but also in the case of phenomenal features.

Finally, I take that there are systematic relations also if we consider a mix of intentional and phenomenal states in a practical argument. One may teach: if the fish stinks like that [experience this smell], throw it away. Then imagine the subject experiences the phenomenal odor of a rotten fish which prompts him to throw it away. However, if the odor is faint, the subject may take time to decide whether to throw the fish away, and this reaction time is systematically linked to the strength of the odor. So, there are systematic relations among phenomenal and intentional states as well as shown by the above *modus ponens*.

If there is systematicity at a high level, the mental, is there systematicity at a low level, the physical? This issue has to be faced by confronting the cotemporally material dependence (CMD) on which Wilson insists. Surely, if one aims at satisfying the PSPC and "realization indifference" as well, one is saying that for each single token S there could be wildly different Ps on which S supervenes. But if we consider the causal relations in which S is involved, and we should consider these because is on these that we assert that there are high-level laws of the sort discussed by special sciences, we may require a sort of systematic counterpart of supervenience: there cannot be systematic variations at a high level without systematic variations at a low level. And this should not be surprising: laws describe systematic relations. Laws in psychophysics, for instance, do exactly this: describe in mathematical terms the stable ratio between the felt sensation and the stimulus causing it.

This ratio determines a difference in the reactions, in the successive expectations, in the latency of the recovery from the stimulations, and so on. In the case of phenomenal features, the variations are embedded in systematic empirical relations.

Now, the more one considers the systematicity of the mental, the more constraints to be placed on the realizers even in case of singular realization. Systematic relations are constraints on realizability. Hence, not all realizers are fit to support all the systematic relations that you discover at the high phenomenal level.

The overall point, then, is that systematicity is a pervasive property of the relations among mental properties such that if you have some systematicity you have a lot of systematicity, and if you have systematicity all the way through, you can't have realization indifference.

4. Defending Premise (ii)

As I have discussed above, a feature being multiply realized is a primary motivation for the weak emergence of such a feature. However, Wilson denies that multiple realizability is a necessary condition for the proper subset condition to be met. Sometimes it looks like it could be a sufficient one:

> while multiple realizability is a good indicator of when a comparatively abstract ontological and causal joint is in place, that there is such a comparatively abstract joint does not hinge on multiple realizability (68).

However (see Ch. 5 on complex systems), Wilson argues that multiple realizability, if not coupled with the satisfaction of PSPC, is not even sufficient for emergence for in many (most) cases candidates for weak emergence are singularly realized. When this single realizability is the case, reductionist have an easy play and it is difficult to make a strong case for weak emergence in these terms. So, what really make the case for weak emergence are those cases in which a feature's powers are a subset of the powers of the realizers on which it cotemporally materially depends, and this may happen to be multiply realized.

As we have seen above, though, many analyses of realization crucially insist on having the weakly emerging features as multiply realized. This is the case with, at least, functional, mereological and determinable-based realization, but there are appeals to multiple realization also in ontologically explanatory realization. I think this appeal is due to the point I was mentioning in the first section: the more a feature or property can figure in type different counterfactuals the more its causal power is well established and robust. So, even in cases in which a feature is singularly realized, more than considering its actual subset of powers, one considers its *counterfactual* subset of powers, those that guarantees distinctive efficacy with respect to the superset powers on which it CMD. It is this the way in which the causal autonomy is robustly vindicated: we can establish the causal autonomy only if a feature makes some difference in a number of significant and causally different scenarios. And the best way to put it is to say that the Ss must determine a "realization indifferent regularity" (cf. Antony and Levine 1997), where the Ps are the differential realizers, no matter whether singularly or multiply. This means that a weakly emergent feature is a "tracker" (82) of difference makers (being causally efficacious) that could determine (potentially, i.e. counterfactually) a realization indifferent regularities (66-69). Such indifference can be as permissive as one can imagine it to be, as per Fodor's (1974) famous schema for special sciences. If S causes S* while CMD on P and P* respectively, this does not amount to S being a new power, for P* may well be caused by a disjunction of low-level properties P1, P2, …, Pn in each case S is instantiated. Suffice that all these Ps have a power in common, the one that satisfy the PSPC via S. I think this is enough to maintain premise (ii). In what follows I will refer to this premise in the shorthand terms: Emergence entails realization indifference.

5. Defending Premise (iii)

The third premise asserts that systematicity entails that realization indifference cannot hold. The following argument runs in support of this premise.

(1) If property S is systematic, the properties logically or empirically related to it are mentioned by or are causally covered by the same or similar laws and regularities.
(2) The Ps on which S cotemporally materially depends (CMD), should follow the same pattern of systematicity shown by S.
(3) If property S is realization indifferent, then it CMD on Ps that are not covered by the same law.
(4) If they are not covered by the same law, the Ps have different projectability patterns and support different counterfactuals.
(5) If they have different projectability patterns and support different counterfactuals, they do not establish the same systematic relations.
(6) If they do not establish the same systematic relation, property S cannot be realization indifferent.

Consider again the case of doubling the intensity of the stimulus. This case rests on using the same predicate, referring to the same property, as being in pain, so using the same projectability, and then embedding that predicate into the same law. But if we want serious realization indifference, this is not allowed, for the pattern of the projectability and counterfactual support of predicates and properties at the high-level disregard the patterns of predicates and properties at the low level. If these patterns are so distant, how can the patterns of projectability and counterfactual support at a high level be the same? These can be the same to a very limited range. For instance, you may realize a lever with iron or with wood to be included in the same machine or in two functionally identical machines: possibly the rigidity of the lever could be the same, but they may differ concerning resistance to fire or oxidation. One may say, this is not relevant. That really depends on the context. For, one may operate with the lever in certain contexts that make their resistance to fire or oxidation relevant, and this cannot be established a priori.

Similarly, if the S is a phenomenal property, it establishes systematic relations to other phenomenal or non-phenomenal properties. Consider seeing a ripe tomato. This produces a phenomenal property of appearance of full red. As such, this property is related to appearances of scarlet or crimson by a similarity relation, which could eventually be subsumed under being a determinate of the same determinable relations to those other shades. Now consider the Ps on which the S in question CMD. If the S is to be realization indifferent and respectful of PSPC, it could well be the case that the Ps on which it supervenes do not match the systematic relations established at the phenomenal level. The subset strategy would apply to them as well. But how far? Up to the point where just the P that happens to be CMD on the S in that token case? That would prevent the subset strategy of its generalization power.

One may wonder why the emergentist should accept premise (2) of this subargument. The emergentist can stress that each "level of reality", whatever that expression designates, is characterized by its laws and hence by its own projectability and counterfactual patterns, contra steps (3) and (4), and these laws could be such to determine different systematic relations or the same relations with a different degree of strength.[6] So, what consequences would bear having different systematic relations, if any at all?

[6] For this point I am indebted to Ivan Cotumaccio and Michele Paolini Paoletti, whom I thank.

According to the subset strategy a property is individuated by the set of its causal powers had by all its instances, hence these should be preserved by all its realizers (which are a subset of the causal power had by the single instances and their realizers). So, this set comprises all the properties that share the causal powers had by the realized property. But the causal powers defining the set do have causal relations to other powers. Say, a rubber band is elastic and green. Elasticity is shared among all elastic entities no matter their color. But elasticity determines fragility in cold conditions. Should we consider this as a condition on other elastic entities? I think we should but suppose we rather think not. Then we may ask a different question: should the elasticity also involve a specific ratio between, say, thickness and length of stretchability? If so, then it could be the case that only a specific realizer fits the bill. But if this is the case, then it seems Kim was right after all: each disjunct has its own merits and the high level is just a measure of our ignorance. Here, the slippery slope on systematicity I was mentioning is reflected in a similar slippery slope on causal powers: once the set is determined, several further causal relations are connected to the causal powers belonging to the set and is very difficult to imagine this being a matter of degree, because fixing the degree of resemblance sounds quite arbitrary.

I think this is a metaphysical point. The identity conditions of a property, what a property is, are determined by its causal relations, sometimes called its causal profile: what causes the property and what the property causes. If such relations are not preserved by its realizers, we can firmly question whether the realizing properties are just a superficial simulation of the property we are considering, mimickers of its behavioral performances in the *specific occasion* at hand, rather than the proper realizer of the high-level property we are considering.

6. Previous Attacks

A different and much more articulated attack on realization indifference comes from Tom Polger and Lawrence Shapiro (2016). Consider, they say, two types of entities A and B which are taken to be of the same kind by taxonomic system S1 and of a different kind by taxonomic system S2. If the factors that lead A and B to be differently classified by S2 are among those that lead them to be commonly classified by S1 and the relevant S2-variation between A and B is distinct from the S1 intra-kind variation between A and B, then we have a real case of multiple realization. However, they continue, no real-life examples come to the rescue. This may seem like an a posteriori argument, open to empirical challenges, though. They confront this argument with possible realizers as well, stressing multiple realiza*bility* rather than multiple realizations, but one may wonder how much their point generalizes.

Also, Paul Thagard (2022) has argued that realization indifference (which he calls "substrate independence") is false. Here is his argument, resting on the assumption that any mental process is an information process:

(1) Real-world information processing depends on energy.
(2) Energy depends on material substrates.
(3) Therefore, information processing depends on material substrates.
(4) Therefore, substrate independence is false.

However, one may defend realization indifference by noting that the kind of difference that energy consumption may make is not relevant to the realization of content.

Another attack comes from Matthew Rellihan (2023). He has argued that realization indifference, which is a basic tenet of functionalism, is a much weaker identity criterion than the one defended by the subset theory of realization. This point is much more relevant than the previous one, being devoted to the strategy at issue in Wilson's book. Realization indifference allows for substituting a causal element for another, provided that it satisfies the same functional role. But these elements may have very different causal powers, and having the same causal powers is required by the subset model. So, such realization indifference is not a guarantee of the sameness of causal power. Consider again the lever of iron and that of wood: they may play the same functional role in their respective machine, but the lever in iron may have a different breaking point from that of wood. So same functional role but different causal power: realization indifference is then to be relativized.

Even if I think this is an effective argument, the reply could be that functional identity has to be all the way down: the two levers must respect the same functional definition in all the relevant aspects. Even if this were the case, it is obscure why we should place such a restrictive constraint. With phenomenal properties, this contextual problem is much deeper.

As I have argued, it is not the external condition that constrain the viability of realization indifference, but systematic relations in which the high-level properties are embedded. After all, these are the properties that determine how a subject feels or what it associates that condition with something else. It is now very difficult to see how this can be guaranteed by realization indifference. Such systematic relations by themselves constrain the realizations allowed.

7. A Different Look at the Whole Argument

An alternative way to put the argument I have been defending so far is the following, which I provide in probabilistic terms:

(1) The higher the similarity in the systematicity of the relations, the lower and less probable that the realizers are wildly realization indifferent;
(2) The lower the probability of realization indifference the higher the probability of having the same realizers;
(3) The higher the probability of having the same realizers, the higher the probability of having the same laws involved;
(4) The higher the probability of the same laws involved the less distinct or causally relevant the Ss involved;
(5) The less distinct and causally relevant the Ss involved, the less their projectability and use in appropriate counterfactuals;
(6) The less their projectability and use in counterfactuals the less the autonomy of the special sciences, *pace* Fodor.

What the argument is saying is that if there is a stable relation between an isolated (not systematic) S property and a P property (these Ss and Ps are kinds) then the S is not realization indifferent, and reduction is viable. If, on the contrary, S is embedded in a pattern of regular and rational relations, hence systematic, then the viability of realization indifference is threatened if not completely undermined. I have used "threatened" and "undermined" because the argument has a

probabilistic nature. So, I admit, it is not a knockdown argument, but one that makes the relation between empirical and logical features evident, and the empirical features, as per scientific practice, point to probability rather than certainty.

On the other hand, if to defend the distinctness and causal relevance of the mental one defends their being nonsystematic, possibly one gains the realization indifference but gets closer to local reductions of the sort advocated by Kim. Now, I agree with Wilson that emergence comes in only two varieties and that the strong one comes with a very high cost that would run against physicalism. If I am right that systematicity puts a serious constraint on the viability of weak emergence, at least the one in which multiple realizability plays a crucial role, it seems that emergence in general has very few hopes to be a viable option in metaphysics.[7]

References

Antony, L. and Levine, J. 1997, "Reduction with Autonomy", *Philosophical Perspective*, 11, 83-105.

Davidson, D. 1970, "Mental Events", in Davidson, D., *Essays on Actions and Events*, Oxford: Clarendon Press 1980.

Davidson D. 1974, "Belief and the Basis of Meaning", in Davidson D., *Inquiry into Truth and Interpretation*, Oxford: Clarendon Press.

Dennett, D. 1971, "Intentional Systems", *Journal of Philosophy*, 68, 87-106.

Fodor, J. 1974, "Special Sciences (or: The Disunity of Science as a Working Hypothesis)", *Synthese*, 28, 2, 97-115.

Fodor, J. 1975, *The Language of Thought*, New York: Crowell.

Fodor, J. 1997, "Special Sciences: Still Autonomous After All These Years", *Noûs*, 31, 149-63.

Humpreys, P. 2016, *Emergence: A Philosophical Account*, Oxford: Oxford University Press.

Johnson, K., Hsiao, S., and Yoshioka, T. 2002, "Neural Coding and the Basic Law of Psychophysics", *Neuroscientist*, 8, 2, 111-21.

Kim, J. 1992, "Multiple Realization and the Metaphysics of Reduction", *Philosophy and Phenomenological Research*, 52, 1, 1-26.

Kim, J. 1998, *Mind in a Physical World*, Cambridge, MA: MIT Press.

Kim, J., 1999, "Making Sense of Emergence", *Philosophical Studies*, 95, 3-36.

Polger, T. and Shapiro, L. 2016, *The Multiple Realization Book*, Oxford: Oxford University Press.

Rellihan, M. 2023, "A Familiar Dilemma for the Subset Theory of Realization", *Analytic Philosophy* 64, 68-90.

Thagard, P. 2022, "Energy Requirements Undermine Substrate Independence and Mind-Body Functionalism", *Philosophy of Science*, 89, 70-88.

Wilson, J. 2015, "Metaphysical Emergence: Weak and Strong", in Bigaj, T. and Wüthrich, C. (eds.), *Metaphysics in Contemporary Physics*, Brill: Leiden, 251-306.

Wilson, J. 2021, *Metaphysical Emergence*, Oxford: Oxford University Press.

[7] For comments on a previous draft, I express my gratitude to Ivan Cotumaccio, Michele Paolini Paoletti, Larry Shapiro, and Jessica Wilson.

Wilson on Metaphysical Emergence

Brian P. McLaughlin

Rutgers University

Abstract

I critically examine Jessica Wilson's views concerning the relationship between Weak emergence and Physicalism and between Strong emergence and Physicalism, and also her defense of libertarian free will in *Metaphysical Emergence* (2021).

Keywords: Metaphysical emergence, Weak emergence, Strong emergence, Physicalism, Fundamental interactions, Free will.

Jessica Wilson's *Metaphysical Emergence* (2021) is a wonderful book. It addresses a wide range of central metaphysical issues from an overarching theoretical perspective. Not only is it must-reading for anyone who works on metaphysical emergence, it contains a wealth of material that should be of interest to anyone who works on physicalism, realization, the metaphysics of complex systems, the metaphysics of ordinary objects, consciousness, mental causation, or free will.

As the title of her book makes evident, Wilson is concerned with metaphysical emergence—metaphysical, rather than merely epistemic emergence. More specifically, she is concerned with whether special science and (scientific and folk) mental kinds, properties, and their instances metaphysically emerge, respectively, from physical kinds, properties, and their instances. A central aim the book is to examine the relationship between that issue and physicalism (15).[1] I'll focus on that aim.

What, then, is physicalism? Wilson takes the core idea of physicalism to be that our world is fundamentally physical.[2] What counts as physical? Wilson appeals to a physics-based conception of the physical, with a caveat in response to Hempel's (1969) famous dilemma (23). The first horn of that dilemma is that if by the physical we mean what is posited by current physics, then, since current physics is incomplete and at least to some extent inaccurate, the claim that our world is fundamentally physical is false. The second horn is that if instead we mean what would be posited by an ideally completed physics that is in fact true of our world, then, since we don't know what such a physics would posit, the

[1] Numerals in parentheses are references to page numbers in the book.
[2] She takes the notion of fundamentality as a primitive (31).

Argumenta 10, 1 (2024): 279—290
ISSN 2465-2334

DOI 10.14275/2465-2334/202418.lau
First published: 30 November 2024

claim that our world is fundamentally physical is largely vacuous. Current physics, for instance, has no need of the hypothesis that there are mental phenomena, but mightn't it turn out to be the case that the physics in fact true of our world does? As Wilson conceives of the physical, it is whatever would be posited by the completed physics in fact true of our world, with the following caveat: A mental feature is not to be counted as a physical feature even if that physics would posit it. She calls this constraint on her physics-based conception of the physical "the no fundamental mentality constraint" (23). She uses it to impose a constraint on physicalism: any doctrine deserving of the name 'physicalism' should be incompatible with the physics in fact true of our world having to posit mental phenomena. She doesn't state a "no fundamental chemical" or a "no fundamental biological" constraint. When discussing physicalism, her attention is typically focused on the place of the mental in nature. I think she would, though, accept such additional constraints. It is clear, for instance, that if the physics in fact true of our world would have to posit entelechies or a fundamental vital force, she would take physicalism to be false (8).

Unlike a term like 'causation', the term 'emergence' is a term of art. Its uses are many and varied both in the philosophical and in the scientific literature.[3] Indeed, they are so diverse that one wonders whether there is even any common core idea. Focusing on metaphysical emergence narrows things down. It is fairly common ground in the philosophical literature at least that whenever there is metaphysical emergence, there is something that emerges and *something else* that it emerges from; that metaphysical emergence is incompatible with reduction; that it always involves emergent properties; and, moreover, that the bearers of emergent properties are complex entities: macro-entities constituted by micro-entities.

Wilson maintains that the core idea of metaphysical emergence is that of dependence with autonomy (1). Emergents are dependent on what they emerge from, yet autonomous from them. She is concerned with emergence from the physical. She calls the kind of dependence that she maintains is required for it, "co-temporal material dependence" (1); and she distinguishes two kinds of autonomy: ontological and causal. She states: "The coupling of co-temporal material dependence with ontological and causal autonomy [...] is most basically definitive of the notion of (metaphysical) emergence" (1). Let's consider, in turn, her notions of ontological autonomy, co-temporal material dependence, and causal autonomy.

What ontological autonomy from the physical comes to is just failure of emergents to be identical with anything physical. Following Wilson in using 'feature' as a blanket term for kinds and properties (including relational properties), if a feature S metaphysically emerges from a physical feature P, then S is not identical with P or any other physical feature. Following her in using 'token feature' as a term for a particular entity's having a feature at a time or throughout an interval of time, if a token feature S emerges from a token physical feature P, then S is not identical with P or any other physical feature token. Further, if a feature S emerges from a physical feature, then any entity that has S is not identical with any physical entity. She takes reduction to require identity claims, and so maintains that metaphysical emergence is incompatible with reduction.

[3] See, for example, the essays in Bedau and Humphreys 2008.

Wilson doesn't explicitly state a definition of 'co-temporal material dependence'. But from her discussion (Ch.1), I take it that she holds that an entity's having a feature S at a time t (what she calls "a token feature S") co-temporally materially depends at t on a configuration of fundamental physical particles having a physical feature P at t (what she calls "a token feature P") just in case at t, the configuration of fundamental particles is coincident with the entity and has a physical feature P that minimally nomologically necessitates S. (Wilson suggests how this could be modified should our world turn out to be gunky (24), but the modification needn't concern us here.) I take it that although a physical feature P must minimally nomologically suffice for S if S emerges from P, P needn't be nomologically necessary for S. Co-temporal material dependence on the physical is compatible with an emergent feature's having multiple physical emergent bases. A token of feature S might emerge from a token of feature P, while a different token of feature S emerges from a token of feature P*, where P and P* are distinct physical features.

Turn to causal autonomy. Wilson holds that emergent features have causal powers: powers to produce certain kinds of effects when an entity has them in certain circumstances. She takes token features, an entity's having a feature at a time or throughout an interval of time, to be the primary *relata* of the causal relation (40). She takes token features to have causal powers too, "token powers" (72). By that I take it she just means that they have causal effects in virtue of being tokens of the features in question and the circumstances in which they are instantiated. She distinguishes two kinds of causal autonomy, and uses the distinction to distinguish two kinds of metaphysical emergence. Her distinction between the two kinds of metaphysical emergence plays a major role throughout the book, so let's turn to it.

Wilson characterizes the two kinds of metaphysical emergence as follows:

> Weak Emergence. What it is for a token feature S to be Weakly metaphysically emergent from token feature P on a given occasion is for it to be the case, on that occasion, (i) that S co-temporally materially depends on P, and (ii) that S has a non-empty proper subset of the token powers had by P (72).
>
> Strong Emergence. What it is for token feature S to be Strongly metaphysically emergent from token feature P on a given occasion is for it to be the case, on that occasion, (i) that S co-temporally materially depends on P, and (ii) that S has at least one token power not identical with any token power of P (120).

The definitions include the same first condition, co-temporal material dependence (explained earlier), but their respective second conditions express different kinds of causal autonomy. In cases of Weak emergence, the token feature S is causally autonomous from the token feature P in that it has a different complete causal profile from the complete causal profile of the token feature P: The token powers of the token feature S (i.e., its effects) are a proper subset of the token powers (the effects) of the token feature P. Thus, every effect of the token feature S is an effect of the token feature P, but the token feature P has effects that the token feature S doesn't have. In cases of Strong emergence, a token feature S has at least one token power (one effect) that is not identical with any token power (any effect) of the token feature P; it does so in virtue of feature S's having a causal power not possessed by P.

I regard Wilson's characterizations of Weak and Strong emergence as entirely stipulative, and so to be judged solely in terms of their theoretical fruits. Of each, we should ask whether there are any instances of the kind of emergence in question, and, if so, what theoretical consequences that has. I'll be concerned with whether there are any instances of the kinds in question, and, if so, the theoretical consequences of that for physicalism, where physicalism is understood to be the thesis that our world is fundamentally physical.

Before turning to those issues, however, I want to first briefly consider other notions of emergence in the literature. Some theorists would deny that causal autonomy, in either of Wilson's two senses, is among the conditions "most basically definitive of the notion of (metaphysical) emergence" (1). They maintain that emergent features can be epiphenomena, and so devoid of causal effects.[4] Let's call that kind of emergence "epiphenomenal metaphysical emergence". One might try to characterize it along Wilson's lines in terms of co-temporal material dependence with the null set of causal powers. Wilson discusses epiphenomenalism (97–101, 140–141). She points out that in the literature, the leading candidates for epiphenomena are the phenomenal or qualitative characters of subjective experiences—their what it is like for the subject aspects—, and argues that they are in fact causally efficacious. I agree with her view that they are causally efficacious. Still, the notion of epiphenomenal metaphysical emergence is coherent; it is an *a posteriori* issue whether there is any. Let it suffice to note, then, that although Wilson sometimes seems to suggest that Weak and Strong emergence are the only two basic kinds of metaphysical emergence, I take it that her considered position is that they are the only basic kinds of metaphysical emergence that we have reason to believe may be found in our world. Of course, epiphenomenal emergentists will disagree even with that weaker claim, but I'll say no more about epiphenomenalism.

As concerns a number of other at least apparently different notions of emergence in the literature, Wilson argues either that they fail to be notions of *metaphysical* emergence or else they in fact involve either Weak or Strong emergence. I recommend in this connection reading her chapter "Complex Systems". It is informative, but it would have benefited from a discussion of the notion of emergence used in solid state physics. That notion is certainly not the notion of Strong emergence in her sense. It would have been instructive to know whether she thinks it involves Weak emergence or instead that it isn't a kind of metaphysical emergence, and why. Be that as it may, I'll now focus just on her notions of Weak and Strong emergence.

Weak and Strong emergence are not so-called because Strong implies Weak but Weak doesn't imply strong. Neither implies the other. They are incompatible: It is impossible for a token feature S to be both Strongly and Weakly emergent from a token feature P, for the simple reason that it can't be the case that the token causal powers of S are a proper subset of the token causal powers P and also the case that S has a token causal power not had by P. Given that they are incompatible, one might wonder why she labels them "Weak emergence" and "Strong

[4] See, for example, Chalmers 1996.

emergence".[5] She doesn't explicitly say, but I take it that she so labels them because she holds that Weak emergence from the physical is weaker than Strong emergence from the physical in the following way: Weak is compatible with physicalism, while Strong is not.

Wilson defends the twofold claim that (a) there is Weak emergence and there may well be Strong emergence, and that (b) while Weak emergence is compatible with physicalism, Strong emergence is incompatible with physicalism. This twofold claim will be my central focus.

Wilson tells us that physicalism is committed to Physical Causal Closure: the thesis that "every lower-level physical effect has a sufficient purely lower-level physical cause" (41). (I take it that the thesis isn't supposed to entail causal determinism. A sufficient cause of an effect must determine the objective probability of the effect, but that can be less than 1 if causal determinism is false.) Weak emergence is compatible with Physical Causal Closure, since the causal powers of the emergent will be a proper subset of the causal powers of its physical base. In contrast, Strong emergence, she tells us, is incompatible with Physical Causal Closure: If there is Strong emergence, then there are at least some lower-level physical effects that do not have any purely physical lower-level sufficient cause (41).

Wilson's formulation of Physical Causal Closure invokes a notion of level, and so presupposes a notion of levels in nature. To be sure, proponents of metaphysical emergence standardly maintain that nature is layered, with higher levels metaphysically emerging from lower levels. Wilson could of course appeal to Weak and Strong metaphysical emergence to characterize two different notions of levels in nature. But the Physical Causal Closure thesis is not supposed to entail that there is metaphysical emergence of even the Weak kind. If, then, the notion of levels invoked in Physical Causal Closure is not to be understood in terms of metaphysical emergence, how should it be understood? What is a level? It is uncontroversial that there are macro-micro levels, but they are just a matter of scale. A proper micro-constituent of a macro-entity will be at a lower level, lower scale, than the macro-entity. But any micro-configuration of physical particles that makes up an entity (at a time) will be at the same scale as that entity (at that time). Systems of particles arranged mountain-wise are at the same scale as mountains, and so not at a different level in the micro-macro sense. So what, then, is a level? Wilson discusses that question (24–30), but doesn't commit to a definitive answer to it since she seems to want to remain neutral on certain issues.

I won't pursue the question of how 'level' should be understood in the Physical Causal Closure thesis. The reason is that I think that Wilson needn't appeal to a notion of levels in order to formulate a physical causal closure thesis that is suitable for her purposes. Given her no fundamental mentality constraint, she could reformulate Physical Causal Closure just as the thesis that every physical effect has a sufficient purely physical cause (one that determines its objective probability). She could then claim that if any mental features are Strongly emergent, that thesis is false, and so physicalism is false since there are *fundamenta* that are not physical. (To address the issue of whether there is chemical or biological

[5] The terms 'weak emergence' and 'strong emergence' get used in the literature, though not in a uniform way. I'm here just concerned with her terms 'Weak emergence' and 'Strong emergence' as she defines them.

Strong emergence, issues she doesn't pursue, one could appeal to a no chemical or no biological constraint on the physics-based conception of the physical.)

Mainly for readability, rather than using 'features' and 'token features', I'll now, for the most part, frame the issues in terms of properties (monadic properties, dyadic ones, etc.), and in terms of states and events as the *relata* of the causal relation. Nothing, I believe, will turn on this shift in terminology. Unless I explicitly indicated otherwise, I'll take states and events to be an entity's having a property at a time or throughout an interval of time, and so what she calls a token feature.

Wilson maintains that Weak emergence is widespread among the special sciences yet compatible with our world being fundamentally physical. Reductive physicalism, she holds, requires that every contingent entity, event, or property be identical, respectively, with some physical entity, event, or property, but that isn't required for our world to be fundamentally physical, and so isn't required for physicalism. A kind of non-reductive physicalism could be true (55–58). She doesn't herself embrace non-reductive physicalism, however, at least not across the board. As I mentioned, she takes there to be reason to believe that there may very well be certain cases of Strong emergence, and so reason to believe that even non-reductive physicalism, as a general doctrine, may very well be false; but of that, more shortly. Let's first look more closely at the relationship between Weak emergence and non-reductive physicalism.

Wilson's notion of Weak emergence requires a modification if Weak emergence across the board is supposed to guarantee non-reductive physicalism. The nomological requirement on Weak emergence is that if a feature S Weakly emerges from a physical feature P, then P is minimally nomologically sufficient for S. That condition is compatible with the law linking S and P being a fundamental law of nature, a law that doesn't hold in virtue of other laws and conditions. The notion of Weak emergence is thus silent about whether the laws linking Weak emergents with their physical bases hold in virtue of physical laws and physical conditions. If S is, for instance, a mental property, the law will be a psychophysical law. The existence of fundamental psychophysical laws is incompatible with physicalism, reductive or non-reductive. If mental properties are distinct from physical properties, and there are fundamental laws in which they figure, then it's not true that our world is fundamentally physical, even if the instances of mental properties don't make a non-redundant causal contribution to the course of physical events (or indeed even if they are epiphenomenal). Mental properties and their instances would be, respectively, fundamental properties and property instances. Since Weak emergence is compatible with fundamental psychophysical laws, it is possible for Weak emergence to hold across the board and yet non-reductive physicalism be false. To avoid this result, the condition of co-temporal material dependence must be amended. It must be amended to include the requirement that the law linking S and P not be a fundamental law of nature; it must be a law that holds in virtue of physical laws and physical conditions.

It should be noted that while this amendment is needed if Weak emergence is to serve the purpose in question, the condition of co-temporal material dependence should not be so amended in the characterization of Strong emergence if Strong emergence is to do the work Wilson intends it to do. A Strong emergentist should hold that laws linking emergents with their physical bases are fundamental laws; and so, not ones that hold in virtue of physical laws and physical conditions. Thus, if Weak and Strong emergence are to do the work that Wilson intends, the

two kinds of emergence require different kinds of co-temporal material dependence, not just different kinds of causal autonomy.

It is fairly common for self-billed non-reductive physicalists to claim that although there are contingent objects, events, and properties that are not physical, they are *realized*, respectively, by physical objects, events, and properties. Realizers are supposed to be more ontologically fundamental than what they realize, thus allowing a kind of non-reductive physicalism. This agreement among non-reductive physicalists is thin, however. 'Realization', like 'emergence', is a term of art. We must be told what's meant by the term. Non-reductive physicalists oblige, but there are a number of non-equivalent relations that get called 'realization' in the literature. As Wilson makes clear, she takes Weak emergence to be realization (vii).[6] She readily acknowledges that there are various notions of realization in the literature, but she seems to hold that they all involve the notion of Weak emergence. She seems to view them as invoked to try to help explain how the kind of causal autonomy required for Weak emergence is implemented. Her view seems to be that if there is realization of any of the kinds in question, then there is Weak emergence.

If, as I've argued, in cases of Weak emergence, the laws linking an emergent with its physical bases must be non-fundamental, it cries out for explanation how it is that such laws hold in virtue of physical laws and physical conditions. Non-reductive physicalists typically want an account of realization that yields such explanations. The role-functionalist notion of realization as causal role occupancy, for instance, yields an explanation of why laws that invoke functional properties hold in virtue of physical laws and physical conditions, and so are not fundamental laws, even though functional properties are not identical with the physical properties that occupy the roles in question. The notion of Weak emergence itself won't yield an explanation of how laws citing Weakly emergent properties hold in virtue of physical laws and conditions.

It is important to note, moreover, that while a role functionalist may hold a view of causation according to which functional states and their physical realizers meet the causal autonomy condition for Weak emergence, a role functional needn't hold such a view. Role functionalists hold that a functional state is a second-order state of being in some state or other that has certain causal effects, and that the first-order states that have those effects realize the functional state. It is open to a role functionalist to maintain that a functional state, a state of being in some state or other that has certain effects, does not itself cause those effects. Its realizers do. That's compatible with functional states figuring in causal explanations of the effects in question.[7] But it is incompatible with Weak emergence.

Weak emergence requires that there be a certain kind of causal overdetermination. As Wilson points out, the kind in question will be different from the familiar kind of causal overdetermination that occurs when, for instance, the shattering of a window is overdetermined by two rock throws (40–46). If one of the rocks throws had not occurred, the window would still have shattered, but not in precisely the manner and at precisely the time in which it in fact shattered. Weakly emergent events, if there are such, don't overdetermine the effects of their

[6] See also Shoemaker's (2009) subset view of realization. Wilson tells us that the subset view of realization was first proposed by Michael Watkins (vii).

[7] For details, see McLaughlin 2006, 2015.

physical bases in that way. The effects of a Weakly emergent event will be precisely the same in manner and time of occurrence as those of a proper subset of the causal effects of its physical event base. Wilson regards this kind of overdetermination as unproblematic, since it is compatible with Physical Causal Closure. It is indeed compatible with Physical Causal Closure. But it cries out for explanation how such overdetermination could occur in our world. We need an explanation of how emergent events can have certain causal effects that their physical base events have, even though those effects would have occurred in precisely the same manner and time even if the emergent event had not occurred.

Whether there is overdetermination of the kind Weak emergence requires, and so whether there is Weak emergence, depends on the answers to questions about the *relata* of the causal relation and about the nature of causation. As Wilson points out (40–44), Jaegwon Kim (1998, 2005) wonders what causal work an emergent state or event could possibly be doing were there such overdetermination, given the causal work done by its physical base. A leading non-reductive physicalist response to Kim's no-work objection is that he is assuming a productive notion of causation, and causation is, rather, a kind of counterfactual dependency (Loewer 2007). Whether this response is available to Wilson depends on some issues about which she is silent. If the entity, feature, or time of a token feature are essential to the token feature, then token features are too fragile to serve as the *relata* of the causal relation on a counterfactual theory of causation.[8] It thus matters whether they are essential to the token feature. Wilson is silent about that.

It is, moreover, uncertain why a non-reductive physicalist would have to appeal to the kind of overdetermination required for Weak emergence. That isn't required if role functionalism counts as a kind of non-reductive physicalism, since, as I've noted, it is at least open to a role functionalist not to countenance the kind of overdetermination Weak emergence requires. It also remains open to a non-reductive physicalist to eschew Wilson's view of the *relata* of causal relations as feature tokens in favor of a coarse grained view of events, and to maintain that every event is identical with some physical event, but deny that special science and mental event types reduce to physical event types.[9] Further, it remains open to a non-reductive physicalists who embraces Wilson's view of the *relata* of causal relations as feature tokens to argue that special science and mental tokens have novel causal powers in a way that is compatible with Physical Causal Closure: They could have novel effects without having novel physical effects. It's been argued, for instance, that special science and mental events will screen off their underlying physical bases from having certain non-physical effects that those special science and mental events have.[10]

Notice that if the kind of view of causation last mentioned is viable, then Strong emergence, as Wilson defines it, isn't incompatible with Physical Causal Closure. A Strongly emergent state or event can have an effect that its physical base doesn't have, yet not have any physical effect that its physical base doesn't have. That's compatible with Physical Causal Closure. Wilson's intent, though, is clearly that Strongly emergent features have novel physical effects, physical effects that lack sufficient purely physical causes (54), so that if there are Strongly

[8] See Lewis 1986.
[9] See, for example, Davidson 1970.
[10] See, for example, Yablo 1992.

emergent features, then Physical Causal Closure is false, and hence physicalism is false. She may be taking it as given that an emergent couldn't have a novel effect (one its physical base doesn't have) without having some or other novel physical effect (one its physical base doesn't have). That may be so, but the issue has certainly not been settled. There is no such consensus about causation. I suggest that rather than getting into the weeds about whether a special science state or event could have novel effects without having novel physical effects, Wilson should modify the definition of Strong emergence so that it explicitly requires that Strongly emergent token features have at least one physical effect that their physical token feature base lacks.

To return to Weak emergence, although Wilson has much of interest to say about non-reductive physicalism and causation, she doesn't say enough to establish that any doctrine deserving of the label "non-reductive physicalism" requires appeal to the kind of overdetermination Weak emergence requires. Moreover, if a non-reductive physicalist maintains there is overdetermination of the kind in question, she owes us an explanation of how it is that there is such overdetermination. The notion of Weak emergence won't help to answer that question. As concerns Weak emergence and non-reductive physicalism, then, my main take away points are that it remains unresolved whether there is overdetermination of the sort Weak emergence requires, and so whether there is Weak emergence, and also whether any doctrine that counts as non-reductive physicalism must appeal to Weak emergence.

Let's turn, finally, to Strong emergence. Wilson claims that libertarian free will requires the Strong emergence of decisions and acts of will, and so is incompatible with Physical Causal Closure, and thus incompatible with physicalism (281). Of course, if there is in fact no such libertarian free will, physicalism faces no such threat. The book's jaw dropper is that Wilson maintains that there is "good reason to think that we have free will of libertarian, Strong emergent variety" (281). She makes a case that we have *prima facie* reason to believe that we have libertarian free will, and that that *prima facie* reason has thus far not been defeated. Her considered position seems to be that we are entitled to believe it until it has been defeated. At one point, though, she says something stronger: "I conclude that there is actual free will of both Weak and Strong varieties" (281). That, however, can't be the best way to state the conclusion she intends. Weak and Strong emergence, you'll recall, are incompatible. If decisions or acts of will are Weakly emergent, then they are not Strongly emergent; and if they're Strongly emergent, then they are not Weakly emergence.

In what remains, I'll focus just on Wilson's claim that decisions and acts of will are Strongly emergent. I'll simply assume, for the sake of argument, that a libertarian notion of free will requires that.

Wilson tells us a novel causal power of a Strongly emergent feature will be a novel fundamental power (54), a power to influence the course of physical events that no physical feature has. Indeed, Strong emergentism, she tells us, "is committed to there being at least one other fundamental force beyond those fundamental forces currently posited" (50) by physics. The force would be a configurational force, a fundamental force, yet one that can be exerted only by complex configurations of particles. As she notes (46-49), in McLaughlin 1992, I claimed that one finds this idea in some of the literature in the British Emergentist tradi-

tion, and that such configurational forces are compatible with Schrödinger's equation, and also that it is an empirical question whether there are such forces. I stand by those claims.

I also claimed in McLaughlin 1992 that I am deeply skeptical about whether there are any fundamental configurational forces, that there seems to be no evidence for their existence, and compelling empirical reason to think there are no such forces. I stand by those claims too. Such forces would involve complex configurations of physical particles participating in fundamental interactions in the physicist's sense of "fundamental interactions". As concerns fundamental interactions in that sense, Wilson says whether there are fundamental configurational interactions is an "open empirical question contingent on as yet unconducted experiments establishing that […] one or more fundamental interactions come into play only under certain comparatively complex circumstances" (283). If, however, that were such fundamental configurational interactions, then current physics would be wrong in a deep way that there is no evidence to believe it is. I'll now elaborate on this point, drawing heavily from a pair of superb articles by the physicist Sean Carroll (2021, 2022). I'll briefly sketch things in broad strokes; for technical details presented in an accessible way, see the Carroll articles.

Quantum field theory includes the Standard Model of particle physics and also gravitation in the weak-field limit of general relativity. It doesn't cover gravitation near black holes; it is silent about the very early universe, about dark matter and dark energy, and also about interactions energies below certain thresholds. Conditions required for its applicability are that gravity is weak and interactions involve energy transfers below a certain threshold. But as Carroll (2021, 2022) points out, human brains and our earthly environment fall well within its scope of applicability.

The key point for present purposes is this: In the field dynamics of quantum field theory, interactions are *local*.[11] They are local in that fields directly interact with other fields only at spacetime points. That is to say, the dynamics of each field at any spacetime point are directly influenced only by the values and derivatives of the other fields at that same point, and not by anything happening elsewhere. That fundamental interactions are local is inextricably baked into the theory. Quantum field theory could, for instance, accommodate new kinds of particles and new kinds of fundamental forces. But the discovery of fundamental configurational interactions would refute the theory. It thus isn't just that quantum field theory doesn't now posit fundamental configurational interactions, it cannot countenance them. Such direct fundamental interactions would involve whole regions of spacetime. That is incompatible with relativity theory.

Quantum field theory has been enormously successful in its regime of applicability, and, as noted, human brains fall well within that regime. The truly enormous empirical support quantum field enjoys soundly defeats any intuitions we might have about there being a fundamental force of will.

Still, to be sure, fundamental configurational interactions can't be ruled out *a priori*. Suppose, then, that current physics has gone very badly wrong indeed, since there are fundamental configurational interactions (relativity theory be damned). Suppose further that acts of will are co-temporally materially dependent on complex neural events, which are in turn co-temporally material dependent on

[11] Entanglement is not local, but it isn't an interaction in the physicist's sense.

events involving astronomically complex micro-configurations of physical particles that participate in fundamental interactions, and so locality fails. Physical particles don't obey the same basic equations when they are in a human brain that they obey when inside a block of ice, even though at some scale human brains fully decompose into physical particles.

Suppose all that is so. Why would it follow that there is libertarian free will? Why would the imagined yet undiscovered fundamental force be a force of will, rather than a fundamental configurational physical force? If acts of will are not identical with the events involving the astronomically complex configurations of particles that (by hypothesis) participate as wholes in such fundamental interactions, but only materially dependent on them, then the question remains whether the acts of will themselves participate in fundamental interactions. Any physical event from which an act of will Strongly emerges will (by definition) nomologically necessitate the act of will, as will any other physical event that nomologically necessitates the physical event in question if nomological necessitation is transitive. Mightn't the acts of will only Weakly emerge from their complex physical base events? Mightn't the acts of will even be epiphenomena, devoid of any effects, and so only be epiphenomenally emergent from those complex physical events? I take it that Wilson's answer to both questions would be "No," but I myself don't see why the answers would be "No". I find it deeply obscure how fundamental configurational interactions, even if there were such, could yield libertarian free will.

Since I've focused mainly on what I take to be some remaining issues for Wilson's view, let me once again express my admiration for *Metaphysical Emergence*. There is much of interest in the book that I haven't even touched on. The book will, I believe, contribute to setting the research agenda on a wide swath of metaphysical issues for years to come.

References

Bedau, M.A. and Humphreys, P., eds., 2008. *Emergence: contemporary readings in Philosophy and Science*. MIT Press.

Carroll, S., 2021. Consciousness and the laws of physics. *Journal of Consciousness Studies*, 28 (9), 16–31.

Carroll, S.M., 2021. The Quantum field theory on which the everyday world supervenes. *In*: O. Shenker, M. Hemmo, S. Iannids and G. Vishine, eds. *Levels of Reality: a scientific and metaphysical investigation*, Jerusalem studies in Philosophy and History of Science. Copenhagen: Springer, 27–46.

Chalmers, D.J., 1996. *The Conscious Mind: in search of a fundamental theory*. New York: Oxford University Press.

Davidson, D., 1970. Mental events. *In*: L. Foster and J.W. Swanson, eds. *Experience and Theory*. Oxford: Clarendon Press, 207–224. Reprinted in Davidson 1980.

Davidson, D., 1980. *Actions and Events*. Oxford: Clarendon Press.

Hempel, C.G., 1969. Reduction: ontological and linguistic facets. *In*: S. Morgenbesser, P. Suppes and M. White, eds. *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel*. New York: St Martin's Press, 179–199.

Kim, J., 1998. *Mind in a Physical World: an essay on the mind-body problem and mental causation*. Cambridge, Mass: MIT Press.

Kim, J., 2005. *Physicalism, or Something Near Enough*. Princeton, N.J: Princeton University Press.

Lewis, D., 1986. Events. *In*: *Philosophical Papers Vol. II*. Oxford: Oxford University Press, 241–269.

Loewer, B., 2007. Mental causation, or something near enough. *In*: B.P. McLaughlin and J. Cohen, eds. *Contemporary Debates in Philosophy of Mind*. Hoboken, NJ: Blackwell Publishing, 243–264.

McLaughlin, B.P., 1992. The rise and fall of British emergentism. *In*: A. Beckermann, H. Flohr and J. Kim, eds. *Emergence or Reduction? Prospects for a Nonreductive Physicalism*. Berlin: De Gruyter, 49–93. Reprinted in Bedau and Humphreys 2008.

McLaughlin, B.P., 2006. Is role functionalism committed to epiphenomenalism? *Journal of Consciousness Studies*, 13 (1–2), 39–66.

McLaughlin, B.P., 2015. Does mental causation require psychophysical identities? *In*: T. Horgan, M. Sabates and D. Sosa, eds. *Qualia and mental causation in a physical world: themes from the philosophy of Jaegwon Kim*. Cambridge: Cambridge University Press, 64–104.

Shoemaker, S., 2009. *Physical Realization*. Oxford: Oxford University Press.

Wilson, J., 2021. *Metaphysical Emergence*. Oxford: Oxford University Press.

Yablo, S., 1992. Mental causation. *The Philosophical Review*, 101 (2), 245–280.

Questioning, Rather Than Solving, the Problem of Higher-Level Causation

Erica Onnis

University of Turin

Abstract

In *Metaphysical Emergence*, Jessica Wilson recognises the problem of higher-level causation as "the most pressing challenge to taking the appearances of emergent structure as genuine" (2021: 39). Then, Wilson states that there are "two and only two strategies of response to this problem" (2021: 40) that lead to Strong and Weak emergence. In this paper, I suggest that there might be an alternative strategy—not opposite, but different in kind—to approach this difficulty. As noticed by Wilson, the problem of higher-level causation was formulated and made central by Jaegwon Kim. However, Kim's arguments were grounded on distinct metaphysical principles—including Alexander's Dictum and its analysis in terms of causal powers. Rather than following Kim's formulation and responding to the problem he raised in his own terms, a different approach may be to question the pertinence of the metaphysical framework in which these arguments were originally grounded. The problem of higher-level causation, in other words, might be less "pressing" if ontological emergence came with a less strict and univocal view of causal novelty and ontological relevance.

Keywords: Emergence, Alexander's Dictum, Causation, Causal powers.

1. The Troubles of the Nonreductionist

Jessica Wilson's *Metaphysical Emergence* (2021) is devoted, as the title suggests, to the analysis of metaphysical forms of emergence. Wilson's focus is on special science macro-entities, whose ontological and causal autonomy are issues close to her heart. She ascribes two features to these entities. First, they depend upon certain complex configurations of fundamental entities, being cotemporally materially composed by them. Second, despite this dependence, special science entities exhibit some ontological and causal autonomy, being "[…] distinct from, and distinctively efficacious with respect to, the micro-configurations upon which they depend" (2021: 2). Special science entities, in short, present both (i) cotemporal material dependence on micro-configurations, and (ii) ontological and causal

Argumenta 10, 1 (2024): 291—302
ISSN 2465-2334

DOI 10.14275/2465-2334/20230.onn
First published: 3 July 2023

autonomy. The coupling of these features provisionally defines metaphysical emergence because (i) and (ii) are real features of the entities at issue.

The compatibility between dependence and autonomy in special science entities, however, is a debated issue. This compatibility problem, indeed, corresponds to a generalisation of the more specific problem of nonreductive materialism highlighted by Jaegwon Kim. This issue arises from embracing both ontological physicalism (the claim that all is physical) and property dualism (the claim that psychological properties belong to a domain which is autonomous and irreducible to the physical one (1989: 32)). The topic that Wilson is addressing is a generalisation of Kim's problem because she is not just interested in *mental* properties and powers, but in a wider range of higher-level entities, such as cells, organs, trees, birds, humans, and so on (2021: 1). The autonomy of these phenomena, however, is under the same threat as the mental properties discussed by Kim, because recognising their autonomy requires solving the so-called "problem of higher-level causation".

The problem was first presented by Kim in 1989, when he argued that no physicalist worthy of the name can be a nonreductionist about psychological phenomena. Kim's analysis proceeds as follows. Nonreductionists accept physicalism. Hence, they accept the so-called "causal closure of the physical", i.e., the assumption that every physical event has a sufficient physical cause. This means that "if we trace the causal ancestry of a physical event, we need never go outside the physical domain" (1989: 43). Consequently, nonreductionists admit that physical events can have only physical causes. However, they reject eliminativism, and are therefore realists about mental properties. This entails that to grant a legitimate existence to mental properties, nonreductionists must find a causal work that is done by mental properties *qua* mental properties (we will soon see why, in Kim's view, it must be so).

Yet nonreductionists already subscribed to the causal closure of the physical, so they seem to come to a dead end: if mental phenomena exert a genuine causal efficacy, then the causal closure of the physical is violated (in addition to the problem of overdetermination, because the effect of a mental cause must have a physical cause as well). If the causal closure is respected, on the contrary, mental phenomena have no genuine causal efficacy and, consequently, no genuine existence. In light of this, Kim concludes that a physicalist has to be either a reductionist or an eliminativist, for she has to reject the distinct autonomy of the mental or the mental *tout court*.

Before turning to Jessica Wilson's presentation of the problem, a relevant remark is in order. Among the premises that lead to the nonreductionists' dead end, Kim briefly mentioned the idea that "to be a mental realist, […] mental properties must be *causal properties*" (1989: 43). Kim fully formulated this principle in a later paper focused again on nonreductionists' troubles with mental causation (2006). Here, Kim asks: "[…] what does the commitment to the reality of mental properties amount to? What is the significance of saying of anything that it is real?" (2006: 436). In Kim's opinion, the answer to these questions is provided by the British Emergentist Samuel Alexander, for whom "To be real is to have causal powers" (ibid.). Kim named this principle "Alexander's Dictum" and its importance within the problem of higher-level causation is evident. If the principle is rejected, entities can have a legitimate existence even without exerting causal efficacy. If the nonreductive physicalist has to give up her nonreductionism,

therefore, it is because of Alexander's Dictum. Let's now turn to Jessica Wilson's formulation and treatment of Kim's problem.

2. The Problem of Higher-Level Causation

As already mentioned, Wilson considers the problem of higher-level causation as "the most pressing challenge to taking the appearances of emergent structures as genuine" (2021: 39). The problem, also known as the overdetermination or the exclusion problem,[1] lies in the apparent impossibility, for a higher-level entity, to be distinctively efficacious in a world where every physical effect is supposed to have an equally physical cause. In this framework, if a non-physical cause is admitted, it follows that the same effect has two sufficient causes, leading to a case of causal overdetermination.

For Wilson, the problem presented by Kim can be exhaustively rephrased starting from six premises. Four of them—*Dependence*, *Reality*, *Efficacy*, and *Distinction* (1-4)—are claims about the nature of higher-level entities; the remaining two—*Physical Causal Closure* and *Non-overdetermination* (5-6)—concern the nature of causation. The premises are the following:

> (1) *Dependence*. Special-science features cotemporally materially depend on lower-level physical features […] in such a way that, at a minimum, the occurrence of a given special-science feature on a given occasion minimally nomologically supervenes on base features on that occasion.
> (2) *Reality*. Both special-science features and their base features are real.
> (3) *Efficacy*. Special-science features are causally efficacious.
> (4) *Distinctness*. Special-science features are distinct from their base features. […]
> (5) *Physical Causal Closure*. Every lower-level physical effect has a sufficient purely lower-level physical cause. […]
> (6) *Non-overdetermination*. Except for cases of the double-rock-throw variety, effects are not causally overdetermined by distinct individually sufficient cotemporal causes (Wilson 2021: 41).

Wilson notices that accepting the dependence, reality, efficacy, and distinctness of special science entities implies the failure of one of the two other premises, and the same can be said about the commitment to the last two premises: if both *Physical Causal Closure* and *Non-overdetermination* are accepted, at least one of the features of special science entities listed above must go.

To solve the problem of higher-level causation there are different strategies, each coinciding with the rejection of one or more premises of the list. In Wilson's opinion, substance dualism rejects *Dependence*, eliminativism *Reality*, epiphenomenalism *Efficacy*, and reductive physicalism *Distinctness*. All these strategies succeed in preserving *Physical Causal Closure* and the *Non-overdetermination* requirement, but they do so by weakening the ontological and causal autonomy of special science entities. Wilson's strategy, conversely, consists in accepting the first four premises about higher-level phenomena, alternatively denying the legitimacy of the other two premises. By doing so, she defines her two schemas for emergence. The rejection of *Physical Causal Closure* leads to Strong Emergence, while that of *Non-overdetermination* leads to Weak Emergence. As we will see in the next paragraph, the first produces a metaphysical position that is not compatible with physicalism, while the

[1] Wilson refers to Kim's (1993) and Merricks' (2003) formulations of the argument.

second allows for a position that is compatible with it. In short, Wilson accepts the structure of Kim's argument, but chooses to reject a different premise than the one chosen by Kim and builds her models of emergence starting from this move.

3. Wilson's Two Schemas for Strong and Weak Emergence

In her book, Wilson poses two key questions. The first is what is emergence, while the second is whether there are real cases of emergence in nature. To answer these questions, while curbing the detrimental effects of the problem of higher-level causality, she designs her two schemas for metaphysical emergence.

The forms of emergence she recognises depend upon the satisfaction of two conditions, the *New Power Condition*, and the *Proper Subset of Powers Condition*. The fulfilment of the first one leads to Strong emergence, while the fulfilment of the second one leads to Weak emergence.

3.1 Strong Emergence

The *New Power Condition* states the following:

> *New Power Condition*: Token feature *S* has, on a given occasion, at least one token power not identical with any token power of the token feature *P* upon which *S* cotemporally materially depends, on that occasion (Wilson 2021: 51).

In this case, to fulfil the condition, it is necessary that the higher-level feature *S* has at least one power that its lower-level base feature *P*, on which *S* materially depends, does not have. If this feature *S* has this new power, then that feature can be considered Strongly metaphysically emergent.

The point to clarify, here, is how the fulfilment of the *New Power Condition* leads to Strong emergence. The answer is that a feature having a new fundamental power cannot (by Leibniz's law) be identical to a feature that does not exert that power. The argument leads, therefore, to the ontological autonomy of the feature at issue. As for its causal autonomy, the argument is much the same. The higher-level feature having a novel power can produce an effect that its base feature cannot because the latter has different powers. Being therefore both ontologically and causally distinct because of the presence of a new power, the feature fulfilling the *New Power Condition* is Strongly metaphysically emergent. In Wilson's words:

> *Strong emergence*: What it is for token feature *S* to be Strongly metaphysically emergent from token feature *P* on a given occasion is for it to be the case, on that occasion, (i) that *S* cotemporally materially depends on *P*, and (ii) that *S* has at least one token power not identical with any token power of *P* (Wilson 2021: 53).

3.2 Weak Emergence

Let's turn to the second schema. The *Proper Subset of Powers Condition* states the following:

> *Proper Subset of Powers Condition*: Token feature *S* has, on a given occasion, a non-empty proper subset of the token powers of the token feature *P* on which *S* cotemporally materially depends, on that occasion (Wilson 2021: 59).

To fulfil the condition, it is necessary that the higher-level feature *S* has a proper subset of the powers possessed by the lower-level base feature *P* on which *S* one materially depends. If the feature at issue has this proper subset of powers, then the feature can be considered Weakly metaphysically emergent.

Similarly to the case of the *New Power Condition*, the fulfilment of the *Proper Subset Condition* entails both ontological and causal distinctness of the higher-level feature. Having different sets of powers, the higher-level and the lower-level features will be ontologically distinct by Leibniz's law and will produce different effects, having causal distinctness due to their different causal profiles (2021: 79). In Wilson's words:

> *Weak emergence*: What it is for token feature *S* to be Weakly metaphysically emergent from token feature *P* on a given occasion is for it to be the case, on that occasion, (i) that *S* cotemporally materially depends on *P*, and (ii) that *S* has a non-empty proper subset of the token powers had by *P* (Wilson 2021: 72).

3.3 How to Be Causally Effective?

As the schemas show, for Wilson it is possible to save the distinctness and causal efficacy of special science entities having (at least) one novel causal power—as in the fulfilment of the *New Power Condition*—or having "a distinctive set (collection, plurality) of powers" (2021: 79)—as in the fulfilment of the *Proper Subset of Powers Condition*. There are therefore two ways in which a higher-level feature—and a special-science entity—can be causally autonomous: it "may have more powers than its base feature", or, alternatively, "fewer powers than its base feature" (2021: 74). If the emergent entity has more powers, some genuine causal novelty appears and violates the Causal Closure. If it has fewer powers, no real causal novelty is involved, but the difference in features and powers had by the entity ensures its ontological and causal autonomy.

In Wilson's opinion, therefore, these are the only two ways in which a higher-level entity can be genuinely efficacious, and for this reason she thinks that every viable account of emergence offered by the literature can be rephrased in her two schemas, which represent the only two appropriate responses to the problem of higher-level causation.

4. Questioning, Rather Than Responding To, the Problem of Higher-Level Causation

In the first paragraph, I described the premises recognised by Kim as underlying the problem of higher-level causation. These are (i) ontological physicalism, (ii) mental realism, and (iii) Alexander's Dictum. These three premises give rise to five of the six premises listed by Wilson. Roughly, *Dependence* and *Physical Causal Closure* originate from ontological physicalism; *Reality* and *Distinctness* descend from mental realism; finally, *Efficacy* derives from the coupling of mental realism with Alexander's Dictum. The sixth premise, *Non-overdetermination*, is independent from the others and is the (unacceptable) consequence, in Kim's opinion, of nonreductionist assumptions. As already suggested, Wilson's and Kim's views about the problem of higher-level causation are structurally similar, even if they solve the problem differently, with Kim rejecting *Distinctness* and Wilson rejecting, alternatively, *Physical Causal Closure* or *Non-overdetermination*.

However, some details of these arguments can be questioned, and in this paper, I would like to focus on those involved with the acceptance of Alexander's Dictum. Specifically, there are three issues that need to be addressed. The first one concerns the Dictum itself: one may want to reject it and assume other criteria about existence. The second one is about the power-based interpretation of the Dictum: one may want to accept the latter, while considering its power-based interpretation as too strict. The third one is about the metaphysical underdetermination of the powers involved in the power-based interpretation: one may want to accept the Dictum and its power-based interpretation, while requiring a differentiation between microscopic physical powers and macroscopic emergent powers. In the next paragraphs, I will examine each of these issues, suggesting that a less strict and univocal view of existence and causal efficacy might render the problem of higher-level causation less "pressing".

4.1 Alexander's Dictum

The first issue is presented here for the sake of the argument, because I think that Alexander's Dictum is reasonable and convincing. I will start with a quick overview about it.

The Dictum is a reformulation of what is known as the Eleatic principle, which owes its name to the visitor coming from Elea who discusses with Theaetetus in Plato's *Sophist* (Oddie 1982). Towards the end of the dialogue, the Eleatic Visitor describes the so-called "battle of gods and giants" (*Soph*. 246e-249d), namely a dispute over the nature of being in which two contrasting views can be recognised. The first one is that assumed by the Gods, i.e., the friends of the forms, who are committed to their immaterial existence; the second, the Giants, are the "earth people", who only grant existence to material and tangible bodies (Assaturian 2021). The Giants' criterion for reality, which can be roughly formulated as "being is being tangible", poses a serious problem: if only tangible bodies exist, how can virtues or souls be accommodated in the resulting ontology? How can something like justice influence the behaviour of the individual, if justice has no tangible body? In this frame, the Eleatic Visitor tries to make the Giants' views more coherent, suggesting that their criterion for reality might be improved. In doing so, he enunciates the Eleatic principle, according to which everything that really is must possess some power or capacity ("τὸ καὶ ὁποιανοῦν τινα κεκτημένον δύναμιν", 246a). The Eleatic principle, therefore, suggests that being, rather than being equivalent to tangibility, is equivalent to having some sort of causal capacity.

Now, the principle (or the Dictum) seems reasonable and convincing because an existing entity unable to produce any sort of causal effects would be hardly conceivable. Still, one might reject it and assume other criteria for existence. Without going too far, while examining free will, Wilson writes that a good reason to take free will at realistic face value is our direct introspective access to it. The fact that we "experience ourselves as seeming to freely choose, in ways transcending any nomological (deterministic or indeterministic) goings-on" (2021: 278) is therefore enough for accepting the genuine existence of free will. Wilson states that "in the absence of good reasons to think that our experience of nomologically transcendent free will cannot be taken at face value, we are entitled to take this experience at realistic face value" (2021: 278). Direct introspective access, therefore, seems a valid criterion for the existence of free will and is different

from Alexander's Dictum, as different as other criteria that have been formulated during the history of philosophy—e.g., being tangible or admitting direct epistemic access, as we already saw, but also being indispensable to our scientific theories (Putnam 1979; Quine 1980), being robust (Levins 1966; Wimsatt 1981 and 1994), and so on. Alexander's Dictum, in short, is not the only reasonable criterion for existence, and admitting other criteria seems to make the problem of higher-level causation less challenging.

4.2 The Power-Based Interpretation of Alexander's Dictum

As mentioned, it is possible and legitimate to assume Alexander's Dictum, namely the principle whereby existence corresponds to the capacity of being causally efficacious. Kim's formulation of the Dictum, however, does not merely equate existence and causal efficacy in general, but rather being with the exertion of causal powers.

This stricter equation might nonetheless be problematic for at least two reasons. The first is historical. As already noticed, Kim states that in Samuel Alexander's opinion being is having some causal powers (2006),[2] but this attribution originated from a misunderstanding. In *Space, Time and Deity* (1920), Alexander expresses an anti-epiphenomenalist position on consciousness, stating that epiphenomenalism is to be rejected (among other reasons) because "it supposes something to exist in nature which has nothing to do, no purpose to serve, a species of noblesse which depends on the work of its inferiors, but is kept for show and might as well, and undoubtedly would in time be abolished" (1920: Vol. II, 8). Kim translates this passage into a power-based vocabulary, but this approach does not reflect Alexander's intentions, as his view of causation was closer to that of Hume than to that of Aristotle. For Alexander, in other terms, causation does not correspond to the exertion of causal powers, but to the relationship of continuity and succession that exist between different regions of Space-Time—the fundamental element of Alexander's metaphysical monism. In *Space, Time and Deity*, Alexander clearly expresses his aversion to the concept of causal power, which, in his view (as also in Hume's), cannot be admitted in our ontologies:

> If all we observe in external events is uniform succession, to impute to one of them a power to produce the other is a fiction, the fiction which Hume set himself to discredit. It may be serviceable anthropomorphism, but it is not science nor philosophy. If there is no power traceable in things, then there is none (1920: 188).[3]

However, Kim is not the only one attributing to British Emergentists some sort of theory of causal powers; Robert McLaughlin did the same in his well-known and

[2] See also Kim: "Prominent […] is the claim that the emergents bring into the world new causal powers of their own, and, in particular, that they have powers to influence and control the direction of the lower-level processes from which they emerge. This is a fundamental tenet of emergentism, not only in the classic emergentism of Samuel Alexander, Lloyd Morgan, and others but also in its various modern versions" (Kim 1999: 5-6).

[3] A little further, Alexander adds: "causality is not the work of power" (1920: 290) and then he goes on to say "The mischief of the conception that a cause has power to produce its effect is that it introduces some mysterious element of connection other than that of simple continuity" (Alexander 1920: 291).

influential paper about the rise and fall of British Emergentism (1992).[4] The problem with these misreadings is that the power-based interpretation, even if only sketched, is not metaphysically neutral (besides being historically inaccurate) and can be misleading.

On the one hand, therefore, the British Emergentists were not committed to a power-based view of emergent causal efficacy. On the other hand, this account of causation might not be the most appropriate for conceptualizing emergence, given its central role in reductionist—i.e., anti-emergentist—strategies. This brings us to the second problem with the power-based interpretation of Alexander's Dictum.

Starting from Kim's causal inheritance principle (1993) and arriving at Elanor Taylor's collapse objection (2015), the notion of causal power has played a pivotal role in strategies aimed at excluding the possibility of higher-level causal efficacy. Kim's causal inheritance principle suggests that higher-level causal efficacy is not genuine, but is derivative from the lower-level by means of the inheritance of lower-level causal powers:

> **Causal Inheritance Principle** (CIP): If mental property M is realized in a system at time t in virtue of physical realization base P, the causal powers of this instance of M are identical with the causal powers of P (Kim 1993: 326).

Taylor's argument (2015), instead, focuses on latent dispositional properties. In her view, higher-level causal efficacy is not genuine because the alleged causal powers of emergent, higher-level phenomena correspond to the dispositional properties belonging to the low-level components on which the emergent phenomena depend. These dispositional properties are latent when the components are in isolation, and their effects become manifest only when they are organised in complex manners: hence the illusion that these properties belong to a higher-level.

What I am suggesting here is that the concept of causal power is central to classic reductionist strategies and seems to already carry anti-emergentist implications. Its introduction into the emergentist debate, moreover, is recent and appears to be related to the recovery of the notion of emergence as an alternative view to contemporary reductionism and physicalism. However, this emergence *vs*. reduction battle is played out within the framework of the latter and draws upon its conceptual repertoire, referring to issues such as realisation, dispositionalism, causal inheritance, and so on. Reading—or re-reading—the emergentist debate in this contemporary key is not necessarily a bad thing, but it is important to recognise that doing so is not metaphysically neutral, nor is it the only approach available.

[4] See McLaughlin (1992: 20): "British emergentism maintains that some special science kinds from each special science can be wholly composed of types of structures of material particles that endow the kinds in question with fundamental causal powers. Subtleties aside, the powers in question emerge' from the types of structures in question". McLaughlin cites C.D. Broad, who indeed uses the term 'power' more than Alexander does. A careful reading of Broad's passages in which the term power is used, however, shows that the term is employed in a non-technical way. Broad, who is referenced by Alexander, similarly believes that causation is a matter of regularity, uniformity, and continuity between spatiotemporal regions (see Broad 1925: 454-56).

There are different interpretations of the Eleatic principle—Samuel Alexander and the British Emergentists provided at least one—and these alternatives seem to make the problem of higher-level causation less challenging.

4.3 The Metaphysical Underdetermination of the Power-Based Interpretation of Alexander's Dictum

While it is perfectly possible to accept both Alexander's Dictum and its power-based interpretation, describing emergent causal efficacy in power-based terms might lead to new problems, rather than solving old ones.

Admitting emergent causal powers seems to naturally raise questions about their nature, namely about what kind of powers they are and whether these emergent powers are different from non-emergent ones.

In the first chapters of *Metaphysical Emergence*, Wilson provides some characterisations of these powers by stating that they are fundamentally novel—this is the reason why Strong emergence is incompatible with physicalism. As for fundamentality, Wilson defines it in primitivist terms: the fundamental is simply what God had to create (2014 and 2021). Wilson adds, however, that a nonfundamental power is a summation or aggregation of already existing lower-level powers (2021: 48), so fundamentality is also defined in terms of compositional basicness: a fundamentally novel power is a non-aggregative power.

Fundamentality, however, does not exhaustively define higher-level causal powers, because microphysical causal powers (those possessed by the emergence base) are fundamental as well. At a first glance, therefore, higher-level causal powers seem to differ from lower-level ones simply by being at a different level.

Further information about these novel powers can be gathered in another passage from *Metaphysical Emergence*. Emergent powers may be intended as grounded in fundamental interactions that are different from physical fundamental interactions (i.e., interactions other than strong and weak interactions, electromagnetism, and gravity) (2021: 133).

These suggestions, however, do not really clarify the nature of these emergent powers, how they act, and how they are exerted by their bearers. Wilson simply states that Strong emergence corresponds to the fulfilment of the condition of having (at least) one novel causal power, but what this power is, is left programmatically undiscussed. For Wilson, that of power is an "operative notion [that is] metaphysically highly neutral" (2021: 32) and "no 'heavyweight' notion of powers or causation need be presupposed" (2021: 33).

Now, the absence of a precise description of emergent powers seems to indicate that there is no relevant difference, in Wilson's view, between lower-level and higher-level causal powers. In other words, it may be reasonable to assume that if there had been a relevant difference, Wilson would have highlighted it.

However, by leaving the power-based interpretation of causal efficacy metaphysically underdetermined and disregarding the hypothesis that emergent causal powers might be relevantly different from low-level ones, two suggestions emerge. First, powers are conceived as a sort of universal and undifferentiated currency for causal processes, regardless of the ontological domain in which they appear. Second, this currency is not "bearer sensitive". Even if emergent properties and entities are different from the properties and entities from which they emerge, the powers of the former are not relevantly different from those of the latter. Here, I use the word "relevant"—or "relevantly"—repeatedly because low-level and

high-level causal powers are obviously different in some way, but the crucial difference I am pointing out is not just any difference, but a difference in kind that might be able to weaken the problem of high-level causation.

By examining the nature of causal powers, for instance, it might be discovered that higher-level powers cannot really collapse, while lower-level ones cannot really emerge. Emergent and non-emergent causal powers, in other words, might simply be non-interchangeable powers of a different kind. Let's try to develop this hypothesis.

Traditional (non-emergent) causal powers are often intended as fundamental, (micro)physical powers. A classic example of these powers is the electron's charge, which is mentioned by several authors involved in the debate (Psillos 2006; Marmodoro 2010 and 2013; Engelhard 2010; Williams 2019) and has peculiar properties that are commonly—though not universally—attributed to powers: being fundamental, essential, intrinsic, intrinsically active, and productive. These features accurately describe many microphysical powers, but macroscopic powers seem more difficult to describe in these terms. Defining the electron's charge as a causal power, in short, seems simpler and more accurate than defining my ability to roller-skate as one.

Emergent causal powers, despite being sometimes intended as ontologically fundamental (Wilson 2021; Barnes 2012), are often conceived as nonfundamental, extrinsic, context-sensitive, and constraining (Thorpe 1974; Mitchell 2012; Gillett 2016; Onnis 2021). These properties appear to be not intrinsically causal but rather determinative in a different (perhaps weaker) sense. Carl Gillett (2016), for instance, defines the causal efficacy of emergent phenomena as a role-shaping, non-productive determination which he dubs "machresis". In his framework, machresis is a "non-powerful" relationship that does not involve the exercise of active and productive causal properties but constrains the already existing contributions of the latter, and in so doing determines reality in "making a difference" to the world. The most striking difference between micropowers and emergent powers would therefore be the intrinsic activity and productivity of the former and the extrinsic non-productive constraining capacities of the latter.

It should be noted that the previous analysis is a preliminary and brief examination of the possible differences between non-emergent and emergent powers. However, it might be useful to engage in a more thorough investigation because powers can easily collapse if they are understood as properties that can be indifferently instantiated at both higher and lower levels. Conversely, differentiating between micropowers and macropowers might make this collapse more difficult. For instance, let's suppose that the macroscopic causal powers exerted by a biological complex system require a biological complex bearer. In that case, a non-biological system or a biological isolated component could not instantiate those macropowers, which would therefore become non-collapsible.

Ultimately, overcoming the metaphysical underdetermination of the power-based view by recognising relevant ontological differences between micropowers and macropowers appears to be another promising approach to making the problem of higher-level causation less challenging.

5. Conclusions

In *Metaphysical Emergence*, Jessica Wilson recognises the problem of higher-level causation as "the most pressing challenge to taking the appearances of emergent

structure as genuine" (2021: 39). As I have attempted to show in this paper, the problem might be less "pressing" if emergence were related to a less strict and univocal view of existence and causal efficacy, and to a more detailed examination of the nature of causal powers.

References

Alexander, S. 1920, *Space, Time, and Deity*, New York: Macmillan.

Assaturian, S. 2021, "What's Eleatic About the Eleatic Principle?", *Archai*, 31, e-03122.

Barnes, E. 2012, "Emergence and Fundamentality", *Mind*, 121, 873-901.

Broad, C.D. 1925, *The Mind and Its Place in Nature*, Brace: Harcourt.

Engelhard, K. 2010, "Categories and the Ontology of Powers", in Marmodoro 2010: 49-65.

Gillett, C. 2016, *Reduction and Emergence in Science and Philosophy*, Cambridge: Cambridge University Press.

Kim, J. 1989, "The Myth of Nonreductive Materialism", *Proceedings and Addresses of the American Philosophical Association*, 63, 3, 31-47.

Kim, J. 1993, "The Non-Reductivist's Troubles with Mental Causation", in Heil, J. and Mele, A.R. (eds.), *Mental Causation*, Oxford: Clarendon Press, 189-210.

Kim, J. 1999, "Making Sense of Emergence", *Philosophical Studies*, 95, 1-2, 3-36.

Kim, J. 2006, "Emergence: Core Ideas and Issues", *Synthese*, 151, 3, 547-59.

Levins, R. 1966, "The Strategy of Model Building in Population Biology", *American Scientist*, 54, 421-31.

Marmodoro, A. (ed.) 2010, "The Metaphysics of Powers: Their Grounding and Their Manifestations", Abingdon: Routledge.

Marmodoro, A. 2013, "Aristotelian Powers at Work: Reciprocity Without Symmetry in Causation", in Jacobs, J.D. (ed.), *Causal Powers*, Oxford: Oxford University Press, 57-76.

McLaughlin, B. 1992, "The Rise and Fall of British Emergentism", in Bedau, M. and Humphreys, P. (eds.) 2008, *Emergence. Contemporary Readings in Philosophy and Science*, Cambridge, MA: MIT Press, 19-59.

Merricks, T. 2003, *Objects and Persons*, Oxford: Clarendon Press.

Mitchell, S. 2012, "Emergence: Logical, Functional and Dynamical", *Synthese*, 185, 171-86.

Oddie, G. 1982, "Armstrong on the Eleatic Principle and Abstract Entities", *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 41, 2, 285-95.

Onnis, E. 2021, *Metafisica dell'emergenza*, Turin: Rosenberg & Sellier.

Psillos, S. 2006, "What Do Powers Do When They Are Not Manifested?", *Philosophy and Phenomenological Research*, 72, 1, 137-56.

Putnam, H. 1979, *Mathematics, Matter and Method*, *Philosophical Papers*, Vol. 1, Cambridge: Cambridge University Press.

Quine, W.V.O. 1980, *From a Logical Point of View: Nine Logico-Philosophical Essays*, Cambridge, MA: Harvard University Press.

Taylor, E. 2015, "Collapsing Emergence", *The Philosophical Quarterly*, 65, 261, 732-53.

Thorpe, W.H. 1974, "Reductionism in Biology", in Ayala, F., Ayala, F.J. and Dobzhansky, T. (eds.), *Studies in the Philosophy of Biology: Reduction and Related Problems*, Berkeley: University of California Press, 109-38.

Williams, N.E. 2019, *The Powers Metaphysic*, Oxford: Oxford University Press.

Wilson, J. 2014, "No Work for a Theory of Grounding", *Inquiry*, 57, 5-6, 535-79.

Wilson, J. 2015, "Metaphysical Emergence: Weak and Strong", in Bigaj, T. and Wüthrich, C. (eds.), *Metaphysics in Contemporary Physics*, Leiden: Brill, 251-306.

Wilson, J. 2021, *Metaphysical Emergence*, Oxford: Oxford University Press.

Wimsatt, W.C. 1981, "Robustness, Reliability, and Overdetermination", in Brewer, M. and Collins, B. (eds.), *Scientific Inquiry and the Social Sciences*, New York: Jossey-Bass, 124-63; repr. in Wimsatt 2007: 43-74.

Wimsatt, W.C. 1994, "The Ontology of Complex Systems: Levels of Organization, Perspectives, and Causal Thickets", *Canadian Journal of Philosophy*, 20, 207-74; repr. in Wimsatt 2007: 193-240.

Wimsatt, W.C. 2007, *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*, Cambridge, MA: Harvard University Press.

Not So Weak Emergence

Michele Paolini Paoletti

University of Macerata

Abstract

In this article, I shall examine Jessica Wilson's schema for weak emergence in connection with two questions: why are only certain proper subsets of the powers borne by lower-level features associated with higher-level, weakly emergent features? Why is a certain proper subset of the powers borne by a given lower-level feature associated with a certain higher-level, weakly emergent feature, and vice versa? I shall consider and criticize four possible answers to these questions, including Wilson's own view. Finally, I shall suggest my own solution, which is based on something akin to grounding categoricalism. I shall also explore some consequences of accepting my view.

Keywords: Emergence, Physicalism, Grounding categoricalism, Powers, Subset account.

1. Introduction

I shall discuss in this contribution Jessica Wilson's schema for weak emergence. I shall show that this schema comes together with two crucial questions. First question: why are only certain proper subsets of the powers borne by lower-level features associated with higher-level, weakly emergent features? Second question: why is a certain proper subset of the powers borne by a given lower-level feature associated with a certain higher-level, weakly emergent feature, and vice versa?

I shall show that answering such questions implies that one rediscusses, *inter alia*, the compatibility between weak emergence and physicalism. In Section 2 I shall briefly introduce Wilson's schema for weak emergence and the two questions I anticipated above. In Section 3 I shall consider three ways of answering (or dissolving) such questions: the suggestion that they ask for explanations of modal facts; primitivism; deflationism about powers. I shall criticize each way. In Section 4 I shall examine and discuss Wilson's own view. Finally, in Section 5, I shall suggest that one should embrace—with respect to higher-level, weakly emergent features and the powers they confer—something akin to grounding categoricalism. I shall also explore some consequences of accepting this view.

Argumenta 10, 1 (2024): 303—312
ISSN 2465-2334

DOI 10.14275/2465-2334/202418.pao
First published: 30 November 2024

2. Weak Emergence and the Two Questions

Jessica Wilson (2021: 72) presents the following schema for weak emergence:

> (WE) a token feature S weakly emerges (on a given occasion) from a token feature P if and only if, on that occasion, (i) S cotemporally materially depends on P and (ii) S has a non-empty proper subset of the token powers had by P.

Token features are particular property-instances. The properties involved in S and P are properties that belong to different levels of the universe. Cotemporal material dependence may be interpreted in different ways, depending on one's favorite theory of ontological dependence. Finally, token powers need *not* be taken as *sui generis* entities, to be distinguished from P, S and their particular instances. For example, on a deflationary view of token powers, the latter may be taken as descriptions of what token features S and P are able to cause in specific circumstances.

In this contribution, I shall dwell on condition (ii). I shall extend the discussion a bit beyond Wilson's original project of providing a schema for weak emergence. And I shall introduce further issues concerning weak emergence and its compatibility with physicalism.

On condition (ii), token feature P has a certain set of token powers associated with it. Assume that this set includes four token powers: p_1, p_2, p_3 and p_4. Following (ii), token feature S has another set of token powers associated with it. Crucially enough, the latter set includes some, but not all of the token powers associated with token feature P. Namely, the set of token powers associated with token feature S is only a proper subset of the set of token powers associated with token feature P. Assume that the set of token powers associated with token feature S includes three token powers: p_1, p_2 and p_3.

This guarantees that, on the one hand, token feature S is *not* endowed with any novel power with respect to the token feature P on which it depends. If token feature P and all of its powers are physical, the weak emergence of S from P is fully compatible with the acceptance of physicalism. Yet, on the other hand, token feature S has a distinctive causal profile with respect to token feature P. Indeed, the distinctive causal profile of S is associated with distinctive laws of nature and distinctive difference-making considerations.

So far, so good. Let me recall the set of powers associated with P, i.e., p_1, p_2, p_3 and p_4. Call this set the "causal role of P". And the proper subset of powers associated with S, i.e., p_1, p_2 and p_3. Call this proper subset the "causal role of S". Three questions arise.

First question: are *all* of the proper subsets of the causal role of P associated with higher-level token features such as S? For example, is there a token feature S_1 associated with p_1 and p_2, another token feature S_2 associated with p_2 and p_3, and so on?

It seems that the answer to this question must be negative. *Not all* of the proper subsets of the causal role of P are associated with higher-level token features. In most cases, *only some* proper subsets are. In our example, only the proper subset including p_1, p_2 and p_3 is associated with a higher-level token feature such as S. Otherwise, we may turn to postulate the existence of higher-level token features that are scientifically irrelevant. Indeed, their distinctive causal profiles/causal roles may be associated with no distinctive law of nature and no distinctive difference-making

consideration. Thus, such higher-level token features would find no place in the best theories of special sciences.

We grant that *only some* proper subsets of the causal role of P are associated with higher-level token features such as S. In our example, only the proper subset including p_1, p_2 and p_3 (i.e., the causal role of S) is associated with a higher-level token feature, i.e., S itself. The next question is: why is the proper subset made of p_1, p_2 and p_3 the only one (in our case) that is associated with *a* higher-level token feature? Namely, why is it the only one that is relevant for the weak emergence of *a* higher-level token feature?

Another question is in order. Even if we concede—*contra hypothesin*—that every proper subset is associated with a higher-level token feature such as S, it seems that the proper subset made of p_1, p_2 and p_3 is the *only one* that is associated *with S*. And it is associated *only with S*. This seems to happen in the actual world not by sheer coincidence, but at least as a matter of nomological necessity. Thus, why is the very proper subset made of p_1, p_2 and p_3 (i.e., the causal role of S) the only one that is associated with S—and only with S? Why is it not associated with any other higher-level token feature? More strongly: why *can't* it be associated—at least as a matter of nomological necessity—with any other higher-level token feature? And why *can't* S have—at least a matter of nomological necessity—any other proper subset of powers associated with it, i.e., any other causal role? In sum, why must S and its causal role be associated with each other (*and only with each other*) at least as a matter of nomological necessity?

We have two questions to face:

1. Why is the proper subset made of p_1, p_2 and p_3 the only one that is associated with a higher-level token feature?
2. Why must S and the proper subset made of p_1, p_2 and p_3 (i.e., its causal role) be associated with each other (and only with each other) at least as a matter of nomological necessity?[1]

3. Three Attempts

These questions cannot be dismissed by claiming that they look for explanations of *modal* facts. First of all, question (1) is not explicitly put in modal terms. Moreover, many questions in the business of metaphysics and philosophy of science are actually put in modal terms, insofar as they ask for explanations of what *can* and *cannot* happen.

Suppose now that, in order to answer both questions, we embrace some sort of *primitivism*. Namely, suppose that we claim that it is a primitive and inexplicable fact of the matter that the proper subset made of p_1, p_2 and p_3 (i.e., the causal role of S) is the only one that is associated with a higher-level token feature. And, more crucially, that that proper subset is only associated with token feature S and S is only associated with that proper subset.

[1] Elder (2004 and 2011) considers similar questions with respect to the restricted composition of everyday objects and with respect to micro-physical causation. In a similar vein, Inman (2018) raises the following problem with respect to the essences of natural substantial kinds: if such essences were nothing but sets of specific properties, why would such properties be unified/clustered together? He criticizes several attempts to solve this problem, e.g., by appealing to homeostatic mechanisms or to specific laws of nature. And, as we shall see, he embraces a non-reductionist solution similar to the one I suggest here.

To make sense of this situation from an ontological standpoint, we may hold that there is some irreducible relation R that links S (and only S) with its causal role (and only with it). Consider now P, i.e., the physical, lower-level token feature. As far as P and its token powers are concerned, R does *not* link any other proper subset of those powers with any other higher-level token feature. Moreover, that R holds between S and its causal role has no further metaphysical explanation. Finally, R may be taken as a nomologically necessitating relation, i.e., as a relation that implies certain nomologically necessary goings-on. This seems to answer both questions.

There are three problems with primitivism. The first problem is that it seems to overpopulate our ontology with many irreducible facts of the matter such as: the fact that R holds between S and the very causal role associated with it.

Secondly, such facts are *not* enough in order to answer question (2). It is *not* enough that R holds between S and its causal role in order to guarantee that S is *only* associated with that role and that role is *only* associated with S. In a given possible world, R may hold between S and its (actual) causal role. But it may *also* hold between S and another causal role. In another possible world, R may *not* hold between S and its (actual) causal role, but between S and another causal role. In sum, there should be something else (a negative fact? A totality fact?) that guarantees that S is *only* associated with its causal role and its causal role is *only* associated with S—both in a given possible world and across possible worlds.

Thirdly and finally, that R holds between S and its causal role is an irreducible fact of the matter. Thus, it is a *fundamental* fact. Moreover, this fact constitutively includes a non-physical token feature such as S. Thus, there are fundamental facts with non-physical token features such as S. The constituents of fundamental facts are fundamental.[2] Therefore, non-physical token features such as S are fundamental.

This conclusion may be hard to swallow for physicalists. True: on one plausible interpretation of physicalism (the one embraced by Wilson 2021), physicalism is only taken to hold that the only powers existing in the (actual) universe are physical powers primarily and non-derivatively borne and exercised by physical entities. Therefore, according to this interpretation, every causal going-on turns out to be exhaustively produced and explained by physical powers. This version of physicalism is fully compatible with there being fundamental facts such as: the fact that R holds between S and its causal role. It is also compatible with S's being a fundamental entity, insofar as S is not endowed with novel powers.

However, that R holds between S and its causal role is *not* a purely physical fact. The former also includes S, which is non-physical. Moreover, that R holds between S and its causal role cannot be fully explained in fully physical terms, since it is a fundamental fact. Thus, that R holds between S and its causal role is at odds with a stronger version of physicalism, according to which everything (at least in the actual universe) is physical or can be fully explained in fully physical terms (i.e., in the end, it entirely depends on the physical and only on the physical).

Invoking deflationism about token powers, causal roles and/or properties does not help either. Assume that "S" is nothing but a scientifically relevant but non-physical predicate and the causal role of S is nothing but a complex description of the nomological regularities connected with "S". In this context, it still

[2] See Sider 2011: 126-32.

makes sense to ask why "S" is associated with *a* description of nomological regularities, why it is associated with *that* description and not with other descriptions, why that description is *only* associated with "S", and so on. From the standpoint of physicalists, the answers to such questions should not (irreducibly) invoke non-physical terms and predicates.

Alternatively, one may hold that causal roles are nothing but complex descriptions of possibly regular behaviors, without the need to invoke non-physical predicates such as "S". Fine. Still, some sets of such descriptions may turn out to *correctly* describe the universe and/or be *useful* when describing the universe. And other sets may turn out to be incorrect and/or useless for such purposes. What accounts for the relevant distinction between correct/useful sets of descriptions and incorrect/useless ones? In order to answer this question, one should find some feature or another in the universe. The alternative would be to adopt a radically anti-realist stance on the bearings of such descriptions. But this would be a non-starter for a project on the metaphysics of emergence. And, more importantly, it would leave something unexplained i.e., the fact that only certain sets of descriptions are correct/useful.

4. Wilson's Physicalist Solution

Wilson (2010; 2021: 177-85) puts forward an account of weak emergence based on degrees of freedom. I cannot enter into detail here. Roughly, the idea is that a weakly emergent entity emerges from its base if, *inter alia*, at least one of the degrees of freedom required to characterize its base is eliminated by imposing certain constraints on the base. Such constraints should be entirely placed at the level of the base. In the end, these constraints must be entirely physical or entirely dependent on the physical.

By eliminating specific degrees of freedom, the powers associated with such degrees are eliminated. Thus, weakly emergent entities turn out to have only a proper subset of the powers associated with their bases.

This mechanism is compatible with the acceptance of physicalism, even in its stronger version. Nevertheless, it is necessary to clarify what one means by "physical constraints". Indeed, by "physical constraints", one may first mean "naturalistically acceptable constraints", i.e., constraints that do *not* involve the existence and/or the action of supernatural entities. This understanding is too weak. For it is compatible with the possibility that some of such constraints are irreducibly non-physical and/or result from the exercise of non-physical powers—even if they still belong to the 'natural world'. For example, some of such constraints may irreducibly belong to the biological level of the universe, so that they still belong to the 'natural world', even if they are not physical.

Secondly, by "physical constraints", one may mean "constraints that necessarily operate through and come together with specific physical processes and changes". This understanding is still too weak. Indeed, if one were to believe in irreducible downward causation, some of such constraints could still be non-physical and/or be caused by irreducibly non-physical entities and/or result from the exercise of non-physical powers—insofar as, in all such cases, the relevant constraints operate through and/or are caused through specific physical processes and changes (by downward causation). For example, an irreducibly biological constraint may still operate through and/or be caused through specific physical processes and changes (by downward causation).

The relevant understanding of "physical constraints" at work here is a stronger one. A physical constraint is one that only involves (in itself and in its own causes) entities and processes that are entirely physical[3] and/or entities and processes that entirely depend on further entities and processes that are entirely physical. This understanding of "physical constraints" makes Wilson's mechanism fully compatible with all versions of physicalism. But it may run into the risk of narrowing down the range of weakly emergent phenomena. Some of such phenomena may result from constraints that—for what we know—do *not* clearly satisfy the third characterization of physical constraints. In other terms, we cannot now assume—and we cannot be now sure—that all of the constraints that contribute to weak emergence are such that they only involve entirely physical entities and processes and/or entities and processes that entirely depend on further entities and processes that are entirely physical.

At any rate, with respect to questions (1) and (2), Wilson's mechanism does *not* provide satisfactory answers. First of all, the characterization of weak emergence in terms of degrees of freedom only provides a *sufficient* condition for weak emergence. Thus, it is *not* guaranteed that every weakly emergent entity will arise through this sort of mechanism. Secondly and more importantly, it seems that *not* every possible elimination of the degrees of freedom required to characterize a base is also able to bring about the causal role of a weakly emergent entity (in our case, of a weakly emergent token instance). On the contrary, it seems that only the elimination of *specific* degrees of freedom—and not others—guarantees this result. Why so? Question (1) is left unanswered.

Thirdly and finally, one must still explain why a certain weakly emergent token feature is only associated with a certain causal role and why the latter is only associated with the former. Question (2) is left unanswered.

In reply to this last worry, one may well embrace a view of token features according to which they are nothing but bundles of token powers. Yet, first, one would then be committed to token powers instead of token features. And, secondly, one would still need to explain why *only certain* bundles of token powers (and not others) seem to 'give rise to' or 'be legitimately describable as' token features.

5. Grounding Categoricalism, or Something Near Enough

In my opinion, the best way to answer questions (1) and (2) consists in embracing something akin to 'grounding categoricalism', i.e., the doctrine according to which the causal roles of categorical properties are somehow grounded on those very properties (see, among others, Tugby 2012, 2021, 2022a, 2022b, Yates 2018, Kimpton-Nye 2021 and Paolini Paoletti 2022).

In Paolini Paoletti 2022, I have defended the following form of grounding categoricalism: by virtue of its own essence, the causal role C of a categorical property P (i) is the causal role *of* P, so that it essentially depends (also) on P, (ii) it depends for its origins on P (i.e., it starts to exist as a causal role thanks to P or thanks to the instantiation of P) and (iii) it depends for its continuing to exist (also) on P (i.e., it continues to exist also or only thanks to P or to the instantiation of P). This entails

[3] An entirely physical entity/process is one that, in principle, can be only characterized (with respect to its essence and with respect to all of its features) in physical terms.

that, as a matter of necessity, the existence of C implies the existence of P: necessarily, C cannot exist without P. And it also entails that, as a matter of necessity, C is the causal role of P and of no other property distinct from P.[4]

By the "essence" of something (be it a property or something else), I mean what that entity non-derivatively is (or could be) in all possible circumstances. Namely, the features to be included in the essence of an entity should *not* derive from other features of that entity and they should necessarily come together with that entity whenever it exists. This view of essences is compatible with the possibility that the essence of an entity is identical with that entity or it is only a description of that entity.

My view is compatible with different conceptions of causal roles. Indeed, causal roles may be nothing but descriptions of regular behaviors.

Please also note that, if one believes that all the (nomologically) possible causal roles exist even if they are not associated with any property, one could modify my view as follows: by virtue of its own essence, the causal role C of a categorical property P (i) is the causal role *of* P, so that it essentially depends (also) on P, and (iv) it (also or only) depends on P for its being a causal role that correctly describes the universe and/or that is 'useful' for the purpose of describing the universe. Indeed, not all the (nomologically) possible causal roles that exist correctly describe the universe and/or are 'useful' for this purpose.

At any rate, if, by virtue of its own essence, the causal role C of a categorical property depends in such-and-such a way on P itself, it seems that C obviously depends on the essence of P, i.e., on what P non-derivatively is (or could be) in all possible circumstances.

We can now apply this view to weakly emergent features and their causal roles.

Roughly, there are three facts to be accounted for: that the proper subset that only includes powers p_1, p_2, p_3 is the causal role of *a* token feature; that it is the causal role *of* token feature S and *only* of token feature S (at least as a matter of nomological necessity); that S *cannot* have any other causal role (at least as a matter of nomological necessity).

The first two facts are easily accounted for by my doctrine. The causal role of a token feature S depends on the property involved in that token feature, i.e., the weakly emergent property in S. It is (also or only) by virtue of the property involved in S that causal powers p_1, p_2 and p_3 are put together so as to constitute the causal role of *a* token feature, so that the relevant causal role starts and continues to exist.

Secondly, it is by virtue of that property that such powers constitute the causal role *of* token feature S, and only of it (or only of token features of that property). And this seems to be part of the essence of the causal role of S[5]. Yet,

[4] I offer a proof of this latter thesis in Paolini Paoletti 2022.

[5] The connection between the weakly emergent property involved in S and the causal role C does not merely hold as a matter of nomological necessity. For there is no possible world with other laws of nature in which C is associated with a property distinct from the one involved in S. C, by virtue of its own essence, is only associated with the property involved in S. This seems reasonable in light of the physicalist commitments of weakly emergentists. Indeed, if C were associated with the property involved in S in one possible circumstance and with some other property in another possible circumstance, then there would be nothing at the level of C (nor at the level of the causal powers included in C) to account for this difference.

my view does *not* entail that powers p_1, p_2 and p_3 turn out to be non-physical. Indeed, such powers may well be physical powers, so that they do *not* depend for what they are on token feature S, nor on the weakly emergent property involved in S. It is only the relevant causal role made of powers p_1, p_2 and p_3 that depends on the weakly emergent property involved in S.

In Paolini Paoletti 2022, I have also defended the following thesis: the categorical property P can have other causal roles different from C in other possible worlds and/or at other times. When applied to weakly emergent properties/token features and the causal roles associated with them, this is at odds with the third fact to be accounted for: that the token feature S (and, presumably, the weakly emergent property involved in it) cannot have any other causal role (at least as a matter of nomological necessity).

If we wish to stick to this fact, we can argue that, as a matter of metaphysical necessity, the weakly emergent property involved in S is realized by causal role C and only by C, so that it cannot have any other causal role. Namely, the weakly emergent property involved in S necessarily depends for its being causally effective on (i.e., is realized by) causal role C and only on it. I assume that dependence for causal effectiveness (i.e., realization) and the other relations of dependence mentioned above are distinct and non-equivalent. I shall expand on this point in a few lines.

Something similar to the solution I suggest here is explored by Wilson (2021: 96-97) in reply to Melnyk (2006). Wilson objects to this solution that scientific truths about scientific features do *not* depend on the presence or on the absence of quiddities (i.e., of qualitative aspects of properties). Moreover, she claims that quiddities are mostly required for transworld individuation, whereas the individuation of properties in worlds that share our laws of nature only proceeds by reference to powers.

What I suggest here is that we *do* need quiddities for metaphysical reasons, i.e., in order to answer questions (1) and (2). Or, at least, we need to appeal to (the essence of) higher-level properties, not fully exhausted by their causal powers. Additionally, not all the facts mentioned in such questions as *explananda* are 'other-wordly' facts. For example, that the proper subset with p_1, p_2 and p_3 is associated with a higher-level token feature is not an 'other-wordly' fact.

In a similar vein and in the footsteps of other authors[6] Inman (2018) suggests that the irreducible essences of higher-level substantial kinds play two roles. First, they structure the modal profiles associated with such kinds, i.e., they connect all the possible ways the relevant substances can be characterized and modified. Secondly, the irreducible essences of higher-level substantial kinds fix the causal profiles associated with such kinds, i.e., all the causal powers the relevant substances possess by necessity whenever they exist.

By embracing my solution, we avoid introducing primitive and *sui generis* connections between token features and proper subsets of powers. However, two problems are left open.

The first problem is that this solution is incompatible with some versions of physicalism. If the causal role of token feature S depends on the higher-level and weakly emergent property involved in S, then it is *not* the case that everything depends on the physical. Secondly, assume that token feature P is physical. P does

[6] Inman (2018: 49) cites Scaltsas (1994: 78-80), Des Chene (1996: 71-75), the Early Modern metaphysician Francisco Suárez (2000), Lowe (2006: 135) and Oderberg (2011).

not depend on the property involved in S. Nor do its physical causal powers depend on that property. However, on the one hand, it seems that the causal role of S depends on the property involved in S. Yet, on the other hand, it seems that the property in S depends—for its being causally effective—on that very causal role. There seems to be a circle of dependence here.

To solve these problems, I suggest that we should first swallow the fact that weak emergence is not so weak. Weak emergence is incompatible with the idea that everything whatsoever is physical or fully depends on the physical.

Moreover, I also suggest that different dependence relations may actually be at stake with the property involved in S and the causal role of S. Indeed, the causal role of S may depend *in a certain respect* (e.g., for its being the causal role of S and for its starting and continuing to exist) on the property involved in S. Yet, the property involved in S may depend *in another respect* (e.g., for its being causally effective, or 'realized') on the causal role of S. Such respects are associated with distinct and non-equivalent dependence relations that may run in opposite directions and still remain by themselves asymmetrical.[7]

By invoking distinct dependence relations, we can then construct distinct and non-equivalent versions of physicalism. We can also generalize in order to make sense of the idea that the physical is more fundamental than the non-physical. Intuitively, we can take into account all the dependence relations that involve physical entities and all those that involve non-physical entities. We can then determine the overall degree of dependence of the former and the overall degree of dependence of the latter. Finally, we can find out that the overall degree of dependence of physical entities is lower than that of non-physical entities, so that the former are more fundamental than the latter.

In sum, there are two lessons to be learnt here. The first lesson is that weak emergence should be accepted in conjunction with metaontological pluralism, i.e., the view that distinct and non-equivalent dependence relations are at stake in the universe. The second lesson is that weak emergence is *not always* compatible with physicalism, i.e., it is not compatible with all forms of physicalism.

It may be objected that my approach is no better than primitivism. Indeed, even primitivism is somehow incompatible with physicalism. And even primitivism turns out to take higher-level, weakly emergent properties as fundamental. However, unlike primitivism, my approach does *not* take the *explanandum* (i.e., the connection between S and its causal role) as a primitive fact of the matter. On the contrary, it explains this connection by appealing to the weakly emergent property involved in S. And my approach postulates no special entity such as the relation R. On the contrary, only the weakly emergent property involved in S and the relevant causal role are taken into account.[8] In turn, the weakly emergent property involved in S is something we are already committed to if we believe that S is a token *feature*. And it need *not* be a universal property. Therefore, *ceteris paribus*, my approach is also ontologically more parsimonious than primitivism.[9]

[7] More on this in Paolini Paoletti 2019 and 2021.

[8] The dependence relations at stake in my approach turn out to be internal relations, i.e., relations whose presence is determined just by the essence and/or the existence of their own relata. On the contrary, the relation R postulated by primitivism is *not* internal. For the weakly emergent property involved in S and its causal role are *not* enough (through their essence and/or existence) to make it the case that R holds between them.

[9] I wish to thank Jessica Wilson and the audience at the Sixth Italian Conference on Analytic Metaphysics and Ontology (L'Aquila 2022).

References

Des Chene, D. 1996, *Physiologia: Natural Philosophy in Late Aristotelian and Cartesian Thought,* Ithaca: Cornell University Press.

Elder, C.L. 2004, *Real Natures and Familiar Objects,* Cambridge, MA: MIT Press.

Elder, C.L. 2011, *Familiar Objects and their Shadows,* Cambridge: Cambridge University Press.

Kimpton-Nye, S. 2021, "Reconsidering the Dispositional Essentialist Canon", *Philosophical Studies*, 178, 3421-41.

Lowe, E.J. 2006, *The Four-Category Ontology: A Metaphysical Foundation for Natural Science,* Oxford: Oxford University Press.

Melnyk, A. 2006, "Realization and the Formulation of Physicalism", *Philosophical Studies*, 131, 127-55.

Oderberg, D. 2011, "Essence and Properties", *Erkenntnis*, 75, 85-111.

Paolini Paoletti, M. 2019, "Respects of Dependence", *Studia Neoaristotelica*, 16, 49-82.

Paolini Paoletti, M. 2021, "Respects of Dependence and Symmetry", *Studia Neoaristotelica*, 18, 31-68.

Paolini Paoletti, M. 2022, "A Brighter Shade of Categoricalism", *Axiomathes*, 32, 1213-42.

Scaltsas, T. 1994, *Substances and Universals in Aristotle's Metaphysics,* Ithaca: Cornell University Press.

Sider, T. 2011, *Writing the Book of the World,* Oxford: Oxford University Press.

Suárez, F. 2000, *On the Formal Cause of Substance: Metaphysical Disputation XV,* translated by J. Kronen and J. Reedy, Milwaukee: Marquette University Press.

Tugby, M. 2012, "Rescuing Dispositionalism from the Ultimate Problem: Reply to Barker and Smart", *Analysis*, 72, 723-31.

Tugby, M. 2020, "Grounding Theories of Powers", *Synthese*, 198, 11187-216.

Tugby, M. 2022a, "Dispositional Realism without Dispositional Essences", *Synthese*, 200, article 222.

Tugby, M. 2022b, *Putting Properties First: A Platonic Metaphysics for Natural Modality*, Oxford: Oxford University Press.

Wilson, J. 2010, "Non-reductive Physicalism and Degrees of Freedom", *British Journal for the Philosophy of Science*, 61, 279-311.

Wilson, J. 2021, *Metaphysical Emergence,* Oxford: Oxford University Press.

Yates, D. 2018, "Inverse Functionalism and the Individuation of Powers", *Synthese*, 195, 4525-50.

Author Meets Critics Session on *Metaphysical Emergence*: Replies

Jessica Wilson

University of Toronto

Introduction

I'd like to start by thanking Simone Gozzano, the patient editor and shepherd of this volume, Massimo Dell'Utri, the benevolent editor-in-chief of Argumenta, and Michele Paolini Paoletti, who initially suggested that an issue of Argumenta be devoted to *Metaphysical Emergence*. Simone, special thanks for your encouragement and your efforts; this is a great honour for me, and you have been a fantastic (and patient) collaborator, in print and in song. I'd also like to sincerely thank my commentators for their illuminating, fruitful, and provocative discussions of my book. The diversity of topics they have addressed, highlighting connections between metaphysical emergence and areas ranging from ontology to property theory to counterfactuals to mereology to quantum field theory to biochemistry and beyond, is truly striking, and a real testament to the wide-ranging import and applications of the notion of metaphysical emergence. Every contribution has given me substantive food for thought. For reasons of space I have focused my replies to each commentator on what I see as the most pressing of their remarks, but of course there is more to say, and I hope and anticipate that these conversations will continue on beyond this volume.

1. Replies to Bellazzi

Bellazzi offers a novel application of Weak emergence as the operative relation between the (broadly biological) function and (broadly chemical) structure of biochemical molecules, such as vitamin B12. As Bellazzi notes, biochemistry stands as a kind of 'hybrid domain' between chemistry and biology, with biochemical kinds understood as having micro-structural features of the sort characteristic of chemical kinds, and certain functions of the sort operative in biological systems. Given that the characterization of a biochemical kind incorporates both structural and functional features, the question arises of how these features stand to one another, as per what Bellazzi calls 'the relation problem'—a problem, and not just a question, reflecting a certain trickiness in identifying a relation capable of

Argumenta 10, 1 (2024): 313—364
ISSN 2465-2334

DOI 10.14275/2465-2334/202419.wis
First published: 30 November 2024

accommodating certain constraints on the connection at issue. These constraints reflect that biochemical kinds are typically both multiply realizable (MR)—such that the same biochemical function can be realised by multiple microstructures—and multiply determinable (MD)—such that the same biochemical structure can realise multiple biochemical functions (see Slater 2009; Bartol 2016; Tahko 2020). These joint features of, or constraints on, the relation at issue are in place for Bellazzi's case study of vitamin B12, whose biochemical functions can be realised by any of four distinct vitamers, and whose biochemical structure(s) can play different roles in human physiology, including in DNA and RNA production, and in hematopoiesis/erythropoiesis.

Bellazzi convincingly argues, to my mind, that taking the relation between biochemical structures and functions to be one of Weak emergence provides an illuminating basis for accommodating MR and MD in the case of vitamin B12, and more generally in other cases of biochemical kinds. I will not repeat the details of her application here, but will rather highlight and discuss what I think are three important ramifications of Bellazzi's discussion for investigations in intra-level metaphysics. I close with some related questions about the specific application at issue.

The first moral of Bellazzi's application is that cases of emergence need not be associated with different 'levels.' Discussions of emergence tend to take for granted that this relation holds between goings-on (in the usual case: features) in different sciences. Hence in my book I focus on cases, e.g., where certain features of ordinary objects of the sort treated by Newtonian mechanics might emerge from features of quantum mechanical aggregates; or where certain thermodynamic properties of complex systems might emerge from properties of statistical mechanical aggregates; or where certain conscious mental states might emerge from neurological and ultimately lower-level physical states; and so on. In the case of biochemical kinds, however, and notwithstanding the connection to chemical and biological kinds and features, what appears to be at issue is the relation between seemingly distinct features of a kind treated by a single special science. The possibility of such intra-level emergence complexifies the structure of special scientific goings-on, both expanding the range of cases which might potentially involve metaphysical emergence, and also suggesting that we should be cautious about assuming that any case of metaphysical emergence is one generating a new 'level' of natural reality.[1] That said, the case of biochemical kinds and features also raises the questions of what relations (most saliently: identity or emergence?) hold between, first, the individual structural and functional components of biochemical kinds, and second, the features in the proximal sciences—i.e., between the structure of a biochemical kind and chemical structure, and the function of a biochemical kind and biological function. I'll return to this issue down the line.

A second moral of Bellazzi's application is that MD is an underappreciated resource so far as theorizing about inter-level metaphysics, and emergence in particular, is concerned. Discussions of emergence often advert to cases of multiple realizability (MR) of a given feature as providing some reason to think that the

[1] A similar moral might be seen as read off of diachronic or 'transformational' conceptions of emergence (see, e.g., Humphreys 1997 and Guay and Sartenaer 2016) as involving fusion or some other interaction at a single level. Bellazzi's moral rather applies to cotemporal emergence of the sort traditionally associated with leveled structure.

feature cannot be treated in reductive (identity-based) terms, and is rather better treated as metaphysically emergent, one way or another. Hence in my book the potential bearing of multiple realizability on a given claim of metaphysical emergence (typically, of the Weak variety) comes up several times. As it happens, a theme of my discussions on this topic is that a feature's being multiply realizable isn't in itself sufficient to establish that the feature is Weakly emergent, at least antecedent to engaging with certain reductionist strategies for accommodating multiple realizability in identity-based terms—most commonly, by taking the lower-level feature to which the higher-level feature is supposed to be identical to be a disjunction of *S*'s realizers; and I also argue that a feature's being multiply realizable isn't necessary for its being Weakly emergent. That said, it remains that the multiple realizability of a higher-level feature is the feature most commonly offered as indicative of a feature's being Weakly emergent. Now, as above biochemical kinds are MR, in that the same biochemical function can be realised by multiple microstructures; but they are also MD, in that a single biochemical structure may realize, or determine, multiple biochemical functions.

To see that MD is an underappreciated resource in theorizing about interlevel metaphysics, note that, notwithstanding that MR poses a prima facie difficulty for reductionism, there is in such cases at least an available candidate lower-level feature (namely, the feature consisting in the disjunction of the multiple lower-level realizers) for the reductionist to appeal to in conformity with their claim that every higher-level feature is in fact identical to some or other lower-level feature. But in cases of MD, it is less clear how an identity-based strategy is supposed to be implemented. Suppose that a single lower-level feature F is capable of determining multiple higher-level features (functional or otherwise) S_1, S_2, and S_3. Each determined feature is, according to the reductionist, identical to some or other lower-level feature, but which one? F can't be identical to just S_1, since in that case F's determination of S_2 and S_3 is unaccounted for. An alternative strategy would be to identify S_1 with some part or aspect of F, and similarly for S_2 and S_3; but even granting that such parts or aspects are available for the identification, as it stands it is unclear that these parts or aspects are properly seen as themselves being lower-level features, as the reductionist requires. Indeed, on some accounts of realization (per, e.g., Shoemaker 2000/2001 and Clapp 2001), token realized features are taken to be proper parts of their realizers. From this perspective, multiple determination poses even more of a challenge to reductionism than multiple realization.

A third moral of Bellazzi's application is that it encodes a distinctive response to the question of which subsets of powers of a given dependence base feature are, or can be, associated with a Weakly emergent feature. In my book, I largely leave it to the scientists to discover which entities and features, and associated powers, are plausibly hypothesized as making sense of natural reality, taking my goal to be that of saying how, given that such-and-such entities and features are supposed to have the key features of metaphysical emergence (as coupling dependence with ontological and causal autonomy), we can make sense of this supposition. I do offer one more specific answer to this question, in the context of discussing an implementation of Weak emergence involving an elimination in degrees of freedom; here the idea is that which degrees of freedom (and associated powers) are eliminated from the characterization of the higher-level feature will often reflect the holding of certain lower-level constraints. But attention to Weak emergence in biochemical kinds provides the basis for a new specific answer to the question

of which subsets of powers are associated with genuine features—namely, that this may be, as Bellazzi puts it, "a product of evolution", and more specifically (as per her forthcoming) that biochemical functions are "associated with a set of chemical powers to bring out a specific effect within biological processes" where these processes are a product of evolution, such that "the relevant chemical powers are indirectly evolutionary selected" (see also Santos et al. 2020). This 'evolutionary' route to identifying which subsets of powers of a given feature are associated with genuine, and moreover Weakly emergent features, is an important part of the background story about why natural reality has the structure it has, which promises to illuminate and apply to kinds and features in biological, ecological, and many other sciences. It also serves to show that there are apparently at least two quite distinct sources capable of generating Weakly emergent features: one broadly synchronic (as in the cotemporal imposition of constraints), and one broadly diachronic. As such, it is unclear whether we should expect a unified metaphysical explanation of which higher-level features come to exist, and why—an important result in its own right.

I want to turn now to raising some questions about Bellazzi's application, falling under the rubric of a single question—namely, how many (potentially instantiated) relations of Weak emergence might be associated with a given biochemical kind?

Let's assume that Bellazzi is right that biochemical functions Weakly emerge from biochemical structures. As above, in being MD, a given biochemical kind may have multiple biochemical functions, each of which would presumably be Weakly emergent from whatever biochemical structure is associated with the kind on a given occasion. So a biochemical kind is plausibly associated with as many Weak emergence relations as the kind has biochemical functions. But now recall that, in being MR, a given biochemical kind may have multiple biochemical structures.[2] And for each such biochemical structure, the question arises of whether it is identical to, or rather (presumably, Weakly) emergent from, a chemical structure. Perhaps each biochemical structure is just identical to some chemical structure, as is suggested by the characterization of biochemistry as "the science that considers the behaviour and effects of chemical processes in biological systems" (Bellazzi, this volume, per Santos et al. 2020). But perhaps there are cases to be made that some or all biochemical structures have only a proper subset of the token powers of associated chemical structures. In that case, a biochemical kind would be associated with as many Weak emergence relations as the kind has distinct realizers. Finally, just as there is a question of what relation holds between chemical and biological structures, there is a question of what relation holds between biochemical and biological functions. Might the latter relation(s) also be ones of Weak emergence? If so, a biochemical kind would be associated with as many Weak emergence relations as the kind has biochemical functions—now running not (as in Bellazzi's case) from biochemical structure to biochemical function, but rather running from biochemical function to biological function.

I offer these questions as further food for theorizing for Bellazzi and others working on the metaphysics of biochemistry. In any case, I'm well convinced that attention to the distinctive characteristics of biochemical kinds points the way towards several new avenues of investigation in the metaphysics of emergence.

[2] I assume that each such structure can serve as a dependence base for any (i.e., all) of the biochemical kind's biochemical functions.

2. Replies to Bennett

In my book, I motivate the powers-based schemas for Weak and Strong metaphysical emergence by attention to the problem of mental/higher-level causation, pressed by Kim (1989 and elsewhere); my basic line is that the two schemas encode the strategies operative in the only responses to Kim which accommodate metaphysical emergence, understood as coupling cotemporal material dependence and (ontological and causal) autonomy. I motivate the schema for Weak emergence, more specifically, by attention to non-reductive physicalist (NRPist) responses to Kim's problematic, which posit diverse relations (functional realization, compositional mechanism, the determinable-determinate relation, and so on) advanced as making sense of how cotemporally dependent higher-level features may be distinct and distinctively efficacious as compared to their physical base features, in a way not involving causal overdetermination of the 'double-rock-throw' variety that makes little sense for the cases at issue. I argue that "a deeper unity of strategy" underlies the seemingly diverse NRPist accounts—namely, that the posited relations[3] each guarantee that, on any given occasion, the higher-level feature has only a proper subset of the token powers of the physical feature upon which it cotemporally materially depends; and I argue that the holding of the Proper Subset of Powers condition, along with the cotemporal material dependence condition, captures what is core and crucial to metaphysical emergence of a physically acceptable variety.

In her contribution, Bennett offers three challenges to this motivation for my account of Weak emergence. The first is that there is an alternative NRPist response to Kim's problematic—Bennett's 'Counterfactual Strategy'—which also encodes "a deeper unity of strategy", but which does not involve any reference to the Proper Subset of Powers Condition. The second is that the Proper Subset Strategy itself does not establish the efficacy of the mental (or Weak emergents more generally). And the third is that the means by which Weak emergent efficacy avoids overdetermination is not as ontologically neutral as I have made it out to be. These challenges are well worth considering; in what follows, I present and respond to each in turn.

2.1 Challenge 1: The Counterfactual Strategy

As noted, I see the deeper unity of strategy underlying diverse NRPist accounts posits as reflecting that their chosen relations guarantee satisfaction of the Proper Subset of Powers condition at the heart of my schema for Weak emergence; but drawing on her 2003 and 2008, Bennett suggests that the underlying unity reflects that the relations posited by NRPists allow implementation of what she calls the 'Counterfactual Strategy' in response to Kim's concerns about overdetermination:

> Talk of overlapping sets of causal powers is not the only way to explain how various intimate relations between the causes defuse the threat of overdetermination. In a (2003) paper, I offered a different explanation. I provided a necessary condition on overdetermination (genuine, 'double-rock' overdetermination), and argued that it is

[3] Not including supervenience or other mere modal correlations, which for various reasons are too weak for physicalist purposes; see Wilson 2005 and McLaughlin and Bennett 2018.

> not met by pairs of causes related in any of the ways [Weak emergentists/NRPists] think that mental and physical phenomena are.
>
> The necessary condition is simply that two causes overdetermine an effect only if had either happened without the other, the effect would still have occurred. That is, causes c_1 and c_2 overdetermine e only if both of the following counterfactuals are nonvacuously true:
>
> $(c_1 \wedge \neg c_2) \rightarrow e$
>
> $(c_2 \wedge \neg c_1) \rightarrow e$
>
> This is a very intuitive test for overdetermination. […] if the test is legitimate, the [Weak emergentist/NRPist] is again in good shape. At least one of these counterfactuals will be vacuous or false when (2003) and only when (2008) the mental and physical causes stand in one of the […] favored relations. […] the basic idea is that on any such relation, the physical base necessitates the weakly emergent mental phenomena, rendering one of the counterfactuals vacuous. (241)

As Bennett's past work makes clear, the necessitation at issue here is metaphysical, such that in every possible world where the physical base feature is instanced, so will be the higher-level mental feature. As such, if c_1 is a mental feature M, and c_2 the mental feature's physical base P, then the counterfactual '$(P \wedge \neg M) \rightarrow e$' will be vacuously true, and the necessary condition for overdetermination will fail to be met.

Bennett offers the Counterfactual Strategy as a kind of 'minimalist' response to Kim's problematic, in the sense that it provides a basis for denying one of the premises in Kim's argument—namely, on Bennett's reconstruction, the premise ('Exclusion') according to which all events that have multiple sufficient causes (that are not themselves causally related) are overdetermined. The Counterfactual Strategy is minimalist in being silent on further details about how, exactly, a higher-level feature might be efficacious in such a way as to avoid overdetermination. That said, as above Bennett does suppose that the Counterfactual Strategy unifies NRPist approaches, and relatedly (as is developed in her 2008) is not available to dualists, including Strong emergentists. In what follows I'll offer three reasons for thinking that the Counterfactual Strategy is subject to problems rendering it unsuccessful even with respect to these minimalist aims. As I'll also observe, the Proper Subset Strategy does not incur these problems, and so is correspondingly advantageous.

2.1.1 Response 1: The Illegitimacy of the Test

Is Bennett's test 'legitimate,' in being a necessary condition on overdetermination, such that failure of one or other counterfactual to be non-vacuously true will get one off the overdetermination hook? No, for it is easy to construct cases of clear overdetermination, where the overdetermining phenomena are nonetheless sensitive to whether the other occurs. Indeed, the whole point of firing squads is to ensure that everyone pulls the trigger, so that no individual is to blame. We can similarly set things up so that Billy and Suzy make a pact that they will each throw the ball at the window only if the other does, so that in the closest worlds where either doesn't throw, neither does the other.

It is an advantage of the Proper Subset Strategy that, unlike the Counterfactual Strategy, it doesn't rely on a condition on overdetermination that is subject to clear counterexample.

2.1.2 Response 2: The Controversy and Context-sensitivity of Counterfactual Assessment

Counterfactual deliberation and assessment are subject to controversy and context-sensitivity. The controversy at issue pertains not so much to the general account of counterfactual truth—most accept some kind of similarity-based account, where a counterfactual is true just in case in the closest world(s) where the antecedent is true, the consequent is true—but rather to the question of how worlds are to be ordered with respect to similarity, given that (as Fine, 1975, nicely established) overall similarity won't do. At present there is no agreement either on more specific criteria of similarity or their ranking. Relatedly, similarity judgements are highly context-sensitive. Bennett briefly registers this in discussing a move according to which (relative to some contexts) events are highly fragile—so fragile that in cases of overdetermination, it turns out to be false that had one but not the other event occurred, then the (same type of) effect would still have been produced.[4] But the more general point is that, given the context-dependence of similarity, whether the counterfactual conditions on overdetermination are or are not met is going to depend on context. Relative to one context, perhaps, there's no overdetermination; relative to another, there is. In that case, Bennett's condition does not provide a clear basis for a response to Kim, but rather pushes the bump in the rug to the question of which contexts are most crucial so far as questions of overdetermination are concerned.

It is an advantage of the Proper Subset Strategy that, unlike the Counterfactual Strategy, it (and the associated response to Kim) isn't subject to the controversy and context-dependence of counterfactuals.

2.1.3 Response 3: failing to distinguish Weak and Strong emergentist responses to Kim

As above, Bennett intends that the Counterfactual Strategy unify Weak emergentist/NRPist responses to Kim's problematic, and distinguish these from anti-physicalist dualist, including Strong emergentist, responses. But as I'll now argue, the Weak and Strong emergentist can implement the Counterfactual Strategy in exactly the same way. Bennett can distinguish these responses, but at the price of taking on board certain controversial metaphysical commitments—commitments not needed to implement the Proper Subset Strategy.

To start, consider the overdetermination counterfactuals for a mental feature M that is supposed to be Weakly emergent. The counterfactual '$(M \wedge \neg P) \rightarrow e$' will likely be non-vacuously true, given the usual assumption that mental states may have diverse physical bases (in a physicalist context: are 'multiply realizable'); for then the nearest antecedent worlds will likely be ones where M has a slightly different physical base (realizer), and M causes e. However (per Bennett's characterization of the NRPist's response to Kim), '$(P \wedge \neg M) \rightarrow e$' will be only vacuously true, given that P metaphysically necessitates M.

[4] Note that this amounts to another 'Counterfactual Strategy' that the NRPist could avail themselves of in response to Kim. Bennett suggests that those endorsing fragile events take the effect to be jointly caused by higher-level and base features, but that diagnosis of the effect's fragility is optional—the fragile event NRPist can just adopt Bennett's minimalist stance and resist calls to provide details about how, exactly, higher-level features enter into causing effects.

Now consider the overdetermination conditionals for a mental feature M that is supposed to be Strongly emergent. The counterfactual '$(M \wedge \neg P) \rightarrow e$' will likely be non-vacuously true, given the usual assumption that mental states may have diverse physical bases (in anti-physicalist context: are 'multiply determined'); for then the nearest antecedent worlds will likely be ones where M has a slightly different physical base, and M causes e. What about '$(P \wedge \neg M) \rightarrow e$?' In her 2008, Bennett argues that the NRPist treatment of this counterfactual "is not available to the dualist": "the dualist cannot say that [this counterfactual] is either false or vacuous […] For the dualist, cases of mental causation do meet the necessary condition on overdetermination". Most relevant here is Bennett's reason for thinking that the dualist (Strong emergentist) cannot claim that the relevant counterfactual is vacuous:

> It is clear that only the physicalist can say that ['$(M \wedge \neg P) \rightarrow e$'] ever comes out vacuous. The dualist cannot, because she does not think that there are any physical events or properties that metaphysically necessitate mental ones. She precisely thinks that there are—at best!—contingent psychophysical laws that link the two. So the dualist denies that there is any legitimate substitute for [P] that would make the antecedent metaphysically impossible. She at most thinks that there are choices of [P] that would make the antecedent nomologically impossible. So the dualist cannot claim that any instance of [the counterfactual] is vacuous. (2008: 290)

This line of thought builds in a controversial metaphysical commitment, however—namely, that Strong emergents are nomologically but not metaphysically necessitated by their physical bases. As I discuss in my (2005), however, there are several views on which Strong emergents are metaphysically necessitated by their physical bases, including a modally consistent Malebranchean occasionalism, a view of properties as essentially constituted by all of the laws into which they enter, and a view of fundamental interactions as holistically unified. Moreover, I argue, the latter two views enjoy considerable empirical support, by contrast with Humean 'anything goes' versions of contingentism which greatly depart from scientific theorizing and practice. Whether or not one accepts any of these views, the fact remains that Bennett's Counterfactual strategy does not itself distinguish between the Weak and Strong emergentist strategies, independent of further controversial assumptions about the modal strength of the connections at issue.

Indeed, upon closer examination even the supposition that the NRPist's favoured relations are such that a physical base metaphysically necessitates a Weak emergent can be denied. Consider functional realization, according to which, e.g., mental feature M is associated with a distinctive causal or functional role, which on a given occasion is played by some lower-level physical feature P. Need P metaphysically necessitate M? Not on causal contingentist views, on which properties and powers may come apart; for on such views there is no guarantee that P, instanced in worlds with different laws of nature, will have the powers requisite unto playing M's causal role. For such a contingentist functionalist NRPist, it might well be that both of the counterfactuals in the Counterfactual Strategy turn out to be non-vacuously true. Correspondingly, the success of the Counterfactual Strategy requires a further metaphysical commitment—namely, the rejection of causal contingentism.

By way of contrast, the Proper Subset Strategy clearly distinguishes between the Weak and Strong emergentist responses to Kim, in a way that is moreover

neutral both on whether either relation holds with metaphysical necessity (requiring only, as per the cotemporal material dependence condition which NRPists and Strong emergentists agree is in place, that emergents supervene with at least nomological necessity on physical goings-on) and on whether causal contingentism is true.

2.2 Challenge 2: No Explanatory Advantage

Putting aside the previous concerns and granting that Bennett's Counterfactual Strategy suffices to undercut the Exclusion premise of Kim's argument, one might wonder whether the Proper Subset Strategy is more explanatory than the Counterfactual Strategy, in going beyond a minimalist response to establish that mental goings-on, in particular, are efficacious. Bennett registers, however, that she is skeptical of this:

> [Wilson's strategy] could solve the exclusion problem and secure the causal efficacy of the mental. But I am still skeptical; I do not think the strategy actually does secure that. All the work is done by Wilson's claim that weakly emergent entities have a nonempty proper subset of the causal powers of their bases. This is the only reason we are guaranteed that weakly emergent entities have causal powers. But Wilson never argues that any particular thing or kind of thing has a non-empty set of causal powers; that is just part of her definition of weak emergence. So those who are inclined to be worried about the causal efficacy of the kinds of phenomena she takes to be weakly emergent—like the mental—will simply deny that they are weakly emergent in her sense. (244)

I agree with Bennett that the Proper Subset Strategy qua response to Kim doesn't itself establish that the mental or any other phenomena is efficacious. The Strategy qua response is at that point in-principle, specifying what it would take for some phenomenon to be Weakly emergent in a way in line with NRPist intentions and accounts. Similarly for the New Power Strategy at the heart of the schema for Strong emergence.

Arguments that mental or other phenomena actually have "a non-empty set of causal powers" come later. Hence after arguing for the in-principle viability of (my conception of) Weak emergence (Ch. 3), I argue that there are good cases to be made that complex systems (Ch. 5), ordinary objects (Ch. 6), qualitative mental states (Ch. 7), and (compatibilist) events of free choosing (Ch. 8) satisfy the conditions in the schema for Weak emergence. I motivate the satisfaction of the Proper Subset of Powers condition by attention to a variety of (empirical, philosophical, introspective, etc.) considerations. In brief (see the chapters for details): for complex systems, satisfaction of the condition mainly hinges on the applicability of the renormalization group method and associated elimination of microphysical degrees of freedom (DOF), coupled with my DOF-based account of Weak emergence; for ordinary objects, satisfaction hinges, alternatively, on the elimination of quantum DOF, on sortal practices of individuation, and on ordinary objects' having metaphysically indeterminate boundaries, understood as per my determinable-based account of metaphysical indeterminacy and coupled with a determinable-based account of Weak emergence; for conscious (qualitative) mental states, satisfaction mainly hinges on perceptions' being determinable, coupled with a determinable-based account of Weak emergence; and for (non-libertarian) free will, satisfaction hinges on an understanding of 'relevant antecedent'

approaches to compatibilist free will according to which the powers of the complex event comprising the relevant antecedents are a proper subset of those associated with the complex event comprising the complete antecedents.

Does the fact that qua response to Kim, the Proper Subset Strategy doesn't itself establish that the mental is actually efficacious mean that the Strategy doesn't have any explanatory advantage over the Counterfactual Strategy? I'm inclined to deny this, for two reasons. First, unlike the Counterfactual Strategy, the Proper Subset Strategy provides an explanatory basis for not just the efficacy, but moreover the distinctive efficacy, of Weak emergents—a distinctive efficacy which tracks difference-making considerations (if my thirst had been differently realized, I would still have reached for the Fresca) associated with comparatively abstract systems of laws or levels of causal grain. Independently of further investigations into which phenomena are actually Weakly emergent, this conception of distinctive efficacy provides the basis for a more compelling NRPist response to Kim than does the Counterfactual Strategy; for it undercuts Kim's incorrect supposition that the distinctive efficacy of a higher-level feature can only lie in the having of a novel power, contra Physical Causal Closure, hence contra Physicalism. Second, unlike the Counterfactual Strategy, the Proper Subset Strategy provides a blueprint for establishing that a given phenomenon is Weakly emergent, and so is not just efficacious but distinctively so—a blueprint that is, as I argue, often realized.

2.3 Challenge 3: Undue Ontological Commitment

Bennett's third challenge is that on the face of it, implementing the Proper Subset Strategy for avoiding overdetermination requires "ontological commitment to trackable, countable causal powers".

> [T]he success of the Proper Subset Strategy entirely depends on the idea that the causal powers of the emergent phenomena are numerically identical to the causal powers of the base. And this in turn requires that token causal powers are the sort of thing that can not only be counted but also individuated. Indeed, it is very, very hard not to imagine them as pebbles in a bucket—and Wilson's diagrams on page 70 suggest that she cannot resist this picture either. But this is a serious and rather discombobulating ontological commitment. I will not argue here that causal powers are not like that, but I suspect others will share my reticence. Even Wilson takes pains to insist that her causal powers are nothing dubious or creepy:
>
> > Talk of powers is simply shorthand for talk of what causal contributions possession of a given feature makes […] to an entity's bringing about an effect, when in certain circumstances… no controversial theses pertaining to the nature of powers, causation, properties, or laws are here presupposed. (32–33; also 45)
>
> But the question is, can she really make good on this neutrality? More precisely, can she assuage my ontological qualms while retaining the nice claim that strictly speaking, there is really only one cause of an effect caused both by a weakly emergent phenomenon and its base? That is the challenge I lay before her. […] My real point here is that one cannot have the Proper Subset Strategy on the cheap; the cost-benefit analysis must be made. We can shoulder the ontological commitment to trackable, countable causal powers and accept the benefits, or we can be squeamish and reject the whole picture. (245–44)

I think this is a fair question, but by way of convincing the skeptic I'm not sure what to say beyond what I've already said. As above, and notwithstanding the convenient schematic representation of powers as 'pebbles in a bucket,' I am explicit about the operative notion of 'power' as simply tracking what (actual or potential) causal contributions the having of a given feature makes when instanced in certain circumstances. As I note by way of proof of metaphysical neutrality, even a contingentist categoricalist Humean can make sense of powers in this sense:

> [E]ven a contingentist categoricalist Humean—someone who thinks that causation is a matter of regularities, such that features have their powers contingently, and that all features are ultimately categorical—can accept powers and the associated notion of causation in the neutral sense(s) here: for such a Humean, to say that an (ultimately categorical) feature has a certain power would be to say that, were a token of the feature to occur in certain circumstances, a certain (contingent) regularity would be instanced. Contemporary Humeans implement more sophisticated variations on this theme; but the point remains that no 'heavyweight' notion of powers or causation need be presupposed in what follows. (33)

So far, so metaphysically neutral. But is it really the case that, as I claim in discussing the schemas, "effectively all participants to the debate can make sense of such identity (non-identity) claims as applied to token (actual or potential) causal contributions (token powers)" (45)? If one has a conception of dispositions or powers as ontological existents, then presumably there is no in-principle difficulty with making sense of these being token identical, in any given case. But as I note by way of proof of metaphysical neutrality, even a contingentist categoricalist Humean can make sense of such identification:

> For example, suppose a contingentist categoricalist Humean wants to take a physicalist approach to the problem of higher-level causation, and so aims (as I will expand on §2.3) to identify every token power of a token higher-level feature with a token power of its lower-level base feature. As previously discussed, such a Humean understands powers in terms of actual or potential instances of a (contingent) regularity. Where the aim is to avoid overdetermination, the Humean may suppose, to start, that the (relevant instances of the) regularities overlap, both with respect to the (single) effect, and with respect to the (single) circumstances in which the two token features occur. If the Humean aims to be a reductive physicalist, they may suppose that such overlap motivates identifying the token features at issue, and hence the associated powers. If the Humean aims to be a nonreductive physicalist, they can reject this identification of features, on difference-making or other grounds of the sort to be discussed §2.3. Such a Humean will suppose that attention to broader patterns of regularities can provide a basis for identifying token powers of token features, even when the token features are not themselves identical. Whether reductive or nonreductive, the contingentist categoricalist Humean can make sense of the claim that some, all, or none of the token powers of token features are identical. As I observed in my (2015: 35), this case is like the case of New York: if we can make it (out) here, we can make it (out) anywhere. (45–6, note 15).

That said, it is worth clarifying that it isn't any part of my view that "there is really only one cause of an effect caused both by a weakly emergent phenomenon and its base"—i.e., the causal power that the mental feature shares with its physical

base. If that were part of my view, I can see why one might be skeptical about the supposed metaphysical neutrality of powers: plausibly, a cause must be some kind of real existent! But on my view it is features (properties, events, etc.) or associated objects which are causes; and talk of powers is (again) just talk of what contribution the having of a given feature can make to the production of certain effects when the feature is instanced in certain circumstances. As such, in any given case of Weak emergence there are indeed two causes on the scene: the two features which share the token power—that is, which are such that their contributions to producing the effect in the circumstances overlap. Relatedly, in her note 6, Bennett says that "given the Proper Subset of Powers strategy, [Wilson] should not think that the effects of mental causes are overdetermined *at all*. For an effect to be overdetermined, it must have at least two distinct causes. But the only sense in which Wilson's [Weak emergentist/NRPist] thinks there are two distinct causes is that there are two distinct phenomena that literally share the efficacious part". Some (e.g., Shoemaker) might want to think of powers or associated potential contributions to the production of effects in mereological terms (as "efficacious parts") but even for such a person, it is the features having the power, not the power itself, that cause the effect. In any case, it's no part of my view that the real 'cause' of a Weak emergent effect is a shared power—so perhaps this clarification will assuage at least some of Bennett's skepticism.

3. Replies to Calosi

In his contribution, Calosi advances a novel mereology—a broadly formal theory of parts and wholes—which aims to (a) accommodate the possibility of metaphysical emergence, without (b) introducing non-mereological structure (as on variations on the theme of hylomorphism; see Koslicki 2008, Fine 2010, and Sattig 2015) or multiplying notions of parthood (as per Cameron 2007 and Canavotto and Giordani 2020). On Calosi's view, a single notion of sum provides the means of accommodating both reducible and irreducible—i.e., emergent—wholes.

The basis of Calosi's mereological framework (following Calosi and Giordani in progress*a* and in progress*b*) is a new conception of sum:

$$Sum(xx, y) \equiv xx < y \wedge$$
$$\forall x\,(\neg x \circ xx \rightarrow \neg x \circ y) \wedge$$
$$\forall x\,(xx < x \rightarrow y < x)$$

Sum is distinct from, and moreover stronger than, existing notions in the literature (see Cotnoir and Varzi 2021), in entailing each other notion while not being entailed by any. The associated mereology assumes an unrestricted composition principle (whereby any plurality of objects composes a Sum), and various axioms governing parthood, including antisymmetry, transitivity, and quasi-supplementation. System in hand, Calosi defines the notions of a 'simple' (having no object as a proper part) and a 'composite,' as the negation of 'simple'; and by appeal to the unrestricted composition principle defines a total function assigning to each object the 'matter' of the object, where the matter of a simple object is the object itself, and the matter of a composite object is the Sum of its components. Calosi is thereby able to distinguish between what he calls a 'Reducible Whole'—a whole that is identical to its matter—and an 'Irreducible Whole'—a whole that is not so identical, which distinction he takes to intuitively correspond to the distinction beween a whole's being 'nothing over and above' its parts (his illustrative cases

being a heap of sand and a lump of clay) vs. 'something over and above' its parts (his illustrative cases being tables, trees, organisms, and statues). And Calosi observes that, given all this, it follows that any simple object is Reducible, and any Irreducible object is composite; but the converse entailments do not hold (some Reducible objects may not be simple; some composite objects may not be Irreducible). Now for the connection to emergence:

> It should be clear why the present proposal has a chance to provide a mereology for emergent wholes: it allows for irreducible wholes that are something over and above their proper parts, i.e., their matter. Indeed, I suggest that, faced with cases of emergent wholes (E) we should endorse the following conditional:
>
> if emergent(x) then Irreducible(x)
>
> That is, Irreducibility as defined above is a necessary condition for emergence. (250–51)

Given that (as above) any Irreducible object is composite, it moreover follows on Calosi's system that

if emergent(x) then composite(x)

That is, being composite is a necessary condition for emergence. Calosi is officially neutral on whether being Irreducible (hence being composite) is sufficient for emergence, since he allows that there might be other 'grounds' for irreducibility. So as I understand Calosi's suggestion, if we have reason to think that some goings-on are emergent, then Calosi's mereology can accommodate them, at least to the extent of satisfying certain key necessary conditions. In this latter respect, Calosi takes his mereology to do better than certain alternative mereologies—most saliently, reductivist conceptions on which composition is identity, and eliminativist conceptions on which there are no composed entities (as per mereological nihilism), which (for reasons that I'll return to below) have been taken to be incompatible with the possibility of emergence, at least of a Strong variety.

By way of further motivating his proposed connection between mereology and emergence, Calosi argues that his account provides a basis for accommodating certain features of emergence as highlighted in my book. First, that emergents depend on yet are distinct from their bases is accommodated in that an Irreducible whole depends on its parts (its 'matter') in that "were we to annihilate its matter, it is unclear that anything would remain of the whole"; yet an Irreducible whole is by definition distinct from its matter. Second, that emergents are typically compositionally flexible is accommodated, at least potentially, in that Irreducible wholes are not identical to compositionally inflexible Sums (Reducible wholes). Third, that emergent entities typically fall under sortals (e.g., 'being a table' or 'being a statue') is accommodated by taking a given sortal to refer to an Irreducible whole as opposed to its Sum (matter). Correspondingly, one need not resort to a non-extensional notion of sum (as on Simons 1987) in order to make sense of, e.g., the applicability of the sortal 'statue' to a lump of clay. Finally, Calosi suggests, his mereological system provides a basis for the leveled structure associated with the special sciences, with special science entities at a level being Irreducible wholes that at each level emerge from sums of Reducible or Irreducible wholes, characteristic of the next level down.

Calosi and Giordani's distinctive mereological framework strikes me as in many ways intuitively plausible and theoretically powerful; in particular, it is a

significant accomplishment to identify and systematize a conception of Sum that unifies and asymmetrically entails existing conceptions. Moreover, I am inclined to agree with Calosi that his application of this framework can be seen as providing a basis for a common characterization of emergent entities as wholes that in some sense exist 'over and above' the mere sum of their parts, which in turn might be seen as confirming an also-common supposition that the notions of emergence and of mereology are deeply connected, such that (at a minimum) emergent entities are necessarily composite, and emergent features are necessarily features of composites.

Even so, in what follows I want to cast a somewhat skeptical eye on the extent to which Calosi's mereology can provide a basis for emergence, and on the more general supposition that emergence and mereology are necessarily connected. I'll start by arguing that while Calosi's application of his mereological framework plausibly provides a basis for *a* conception of emergence, this conception is different both from that which he seemed to have in mind in offering his illustrative cases of Reducible and Irreducible wholes, and from that which I aim to characterize in my book; and I'll draw out certain implications of this result for his project. I'll then highlight some considerations which indicate that the connection between emergence and mereology is not as deep (or necessary) as has sometimes been assumed.

To begin: recall that Calosi characterizes Reducible wholes as those which are (as he puts it) intuitively 'nothing over and above' their parts, with his examples being of unstructured entities or aggregates such as heaps and lumps of clay, and Irreducible wholes as those which are intuitively 'something over and above' their parts, with his examples being those of structured entities such as tables, trees, organisms, and statues; and he wants to make use of the distinction between Irreducible and Reducible wholes to at least make room for entities to be emergent, or not.[5] Now, an initial problem here, which poses a problem for identifying a purportedly Reducible heap or lump of clay with its 'matter,' is that heaps and lumps aren't identical to the sum of their scattered parts, in which case Calosi's mereology deems heaps and lumps Irreducible as opposed to Reducible wholes, and so doesn't distinguish his illustrative paradigm cases (which in turn were supposed to be candidates for non-emergent vs. emergent wholes). In any case, at best the Reducible/Irreducible distinction operative here is apt for distinguishing completely unstructured objects—mere collections, as fusions—from any at-all-structured objects.

Now, the distinction between structured and (completely) unstructured entities is no doubt important. It has played an important role, in particular, in discussions of the metaphysics of ordinary objects, as entities which are structured as opposed to unstructured collections of parts, as in Koslicki's (2008) motivating case of a (structured) motorcycle and an (unstructured) heap of motorcycle parts. But this distinction has not played an important role in debates about whether seemingly higher-level (ordinary, special scientific) goings-on are reducible or rather emergent.

[5] Calosi does not specify whether the emergence at issue is to be understood in Weak or Strong terms. In discussing the application of his framework to accommodating leveled structure of the sciences he seems to have Weak emergence in mind; but on the other hand concerns about whether emergence is compatible with reductive or eliminativist conceptions of composition typically suppose that the emergence at issue is Strong. In any case, which form of emergence is at issue won't matter for my present point.

To see why this is so, note that reductive physicalists, who think that any given special science entity or feature is type identical to some or other (perhaps logically or otherwise complex) lower-level physical feature, *take for granted* that the entities to which special science entities are identical are structurally complex (which is not to say that they are committed to composites as distinct from pluralities, about which more anon). It's no part of the reductive physicalist's view to maintain that tables, trees, organisms, or statues are identical to unstructured entities or aggregates. Rather, to take a toy example, a reductive physicalist might identify a certain table with a relational aggregate of atoms standing in atomic relations (or a disjunction of such aggregates, to allow for the table to persist through some change), and so on.[6] So the distinction between something that is in some sense just an unstructured sum of parts and something that is rather in some sense a structured aggregate isn't, at least in the usual cases, what is at issue in the physicalism debates, or in the related debates over whether or not there are multiple 'levels' of natural reality. And nor is it what is at issue in my attempts (in my book and elsewhere) to characterize metaphysical emergence in a way making sense of the appearances of higher-level reality. Rather, what is at issue in these contexts is the question of whether, *in addition* to whatever massively complex, typically highly structured, lower-level physical goings-on there might be, there are moreover any goings-on which are properly seen as (cotemporally materially) dependent on and (ontologically and causally) autonomous from the (massively complex, typically highly structured) lower-level physical goings-on that emergentists and non-emergentists alike agree exist.

Again, this is not to deny that there might be a different, weaker conception of metaphysical emergence that the broad distinction between unstructured sums and structured wholes might latch onto. It would serve, for example, to characterize an extreme form of reductive physicalist—call them 'the reductive pluralist'—who maintains that every apparently structured entity is really identical to some unstructured lower-level physical entity (or logical construction thereof). My point here is just that Calosi's conception of emergence as 'mirrored in' the distinction between (unstructured) Reducible and (structured) Irreducible entities is not obviously suited to accommodating metaphysical emergence of the sort at issue in debates over leveled structure, and which I aim to characterize.

The previous result has certain implications for Calosi's advertised characterization of his mereology as able to accommodate emergence without requiring additional (e.g., hylomorphic) non-mereological resources or multiplying notions of parthood. For insofar as the conception of emergence for which Calosi's system provides a basis is too weak to distinguish between non-emergent structured entities (of the sort the reductive physicalist accepts) and emergent structured entities (of the sort that Weak and Strong emergentists accept), it remains open that properly accommodating metaphysical emergence might require such additional resources or notions of parthood, after all. That said, it remains unclear to me whether we should be asking our mereological systems to do this work. So far as

[6] Nor is the reductive physicalist's characteristic rejection of there being multiple 'levels' of natural reality (as per, e.g., Heil 2003) based in the supposition that there are no structured wholes. Rather, reductionists as well as emergentists will accept that there are 'levels' of the sort that Calosi offers as 'mirroring' the Weak emergentist conception of multiple levels—though they will then deny that these mereologically-generated levels are tracking what is at issue between them.

I can tell, the conditions I provided on metaphysical emergence in my book don't rely, even indirectly, on any mereological notions.[7]

This brings me to my next topic, which pertains to the question of whether the notions of emergence and mereology are necessarily connected, as in Calosi's claims that if an object is emergent, then it is Irreducible, and (coupled with his supposition that if an object is Irreducible, then it is composite) that if an object is emergent, then it is composite.

Now, it is indeed sometimes claimed that composition is a necessary condition on emergence. For example, Baron (2019) says, "[m]ereological composition is usually thought to be at least a necessary condition on dependence: the emergent entity is composed of the entities from which it emerges" (2210). Calosi (2016a) agrees, saying that "An emergent property is a property which is exemplified by a composite object" (441).

As I see it, however, there are two good reasons to deny that composition is a necessary condition on emergence. First, even if it is granted that an emergent entity must cotemporally depend on a composite entity, as has often (though not universally; see below) been assumed for cases of both Weak and Strong emergence, the bearer of the emergent feature might not be composite. Consider the case of persons and their bodies. It is commonly maintained that persons are emergent, either Weakly or Strongly, in having Weakly or Strongly emergent mental states. But this much doesn't require that persons *themselves* be composite: perhaps they cotemporally depend on composites (bodies, or lower-level aggregates) without themselves having parts.[8] So there can be uncomposed emergent entities, and emergent features (e.g., mental states of non-composite persons) not exemplified by composites. Second, it's unclear that an emergent entity or feature has even to cotemporally depend on anything composite. One sort of possibility here involves a simple entity emerging, Weakly or Strongly, from another simple entity, when the latter is in appropriate circumstances. In the Weak case: perhaps the emergent is a determinable of a more determinate simple entity.[9] Perhaps that's a non-standard case, but it seems coherent to me. Another and quite standard option would involve the emergent entity (feature) cotemporally depending on a plurality (feature of the plurality). In my book I register this possibility, and more generally make room for the base-level goings-on to be pluralities or features of such pluralities (as opposed to, e.g., relational aggregates and features of such aggregates). In any of these cases, there might be emergence of either Weak or Strong varieties in the absence of composition as involving anything like a 'whole.'

[7] Of course, in some cases a given implementation of either Weak or Strong emergence might well involve the supposition that the emergent entities (features) at issue are composed (are features of composed entities); my own degrees-of-freedom-based account of Weak emergence is a case-in-point. But even here, the appeal to mereology is mainly serving as a way of ensuring that the condition on cotemporal metaphysical dependence (encoding the supposition of substance monism generally operative in accounts of emergence), is met; it is not itself serving as the basis for emergent autonomy.

[8] Would persons then not be 'concrete?' I don't see why not, given that they exist in spacetime (see Armstrong 1978).

[9] Note that the determinable-determinate relation is typically not cashed in mereological or related (e.g., conjunctive) terms. Most saliently, to be determinate is not to have a determinable as a proper part: determinates (unlike wholes) do not satisfy anything corresponding to supplementation.

The previous considerations undercut Calosi's necessary conditions on emergence, and more generally suggest that the connection between emergence and mereology might not be as intimate as Calosi and some others have taken it to be.

But what about arguments aiming to show that (the possibility of) emergence is incompatible with reductionist approaches to composition such as composition as identity (CAI), according to which mereological fusions are just identical to the plurality of their parts (see McDaniel 2008, Schaffer 2010, Calosi 2016a and 2016b)? Don't such arguments show that there is a deep connection between emergence and mereology, after all? Though I cannot address all such arguments here, I believe that their conclusions can be resisted, for reasons set out in Bohn 2009 (see also Cornell 2017 for a similar strategy). Bohn focuses his attention on the argument in McDaniel 2008, which Bohn schematically characterizes as follows:

1. Emergent properties are possible
2. If CAI is true, emergent properties are impossible
3. CAI is false

Here the focus is more specifically on Strongly emergent properties. Granting that Strongly emergent properties are possible (a claim with which I agree), why think that such properties would be incompatible with CAI? McDaniel's line of thought is that such an assumption leads to a violation of Leibniz's Law, according to which identicals are indiscernible. To start, let some xx be a plurality of two or more things, and let f(xx) be their compositional fusion. Now, assume that the fusion f(xx) has some Strongly emergent property *F*, understood (by McDaniel) as fundamentally novel as compared to the intrinsic properties of and spatiotemporal relations between the xx.[10] McDaniel then argues that insofar as *F* is fundamentality novel as compared to the intrinsic properties and spatiotemporal interrelations of the xx, *F* can be attributed to f(xx) but not the xx—but in that case, identifying the xx and f(xx) as per CAI would violate Leibniz's Law.

As Bohn correctly notes, however, McDaniel's reasoning here fails to appreciate that there's no problem with taking the plurality xx to have a fundamental collective property. As Bohn puts it, "according to the composition as identity theorist, any emergent property of the fusion should simply be thought of as a terminological variant of a fundamental plural collective property of all the parts, and vice versa. In that way the composition as identity theorist can hold that emergent properties do not violate the principle of indiscernibility of identicals" (221). This seems right to me, and I also agree with Bohn that a similar reply is available in response to those (including Calosi, who in his 2016b argues that a version of CAI is equivalent to mereological nihilism) maintaining that mereological nihilism is incompatible with Strong emergence.

So as it stands I remain unconvinced that emergence of any variety requires that there be composed wholes of the sort that CAI denies exist, or indeed any wholes at all. Pluralities, and even a single object, will do.

All this said, I suspect that there is new work for Calosi's mereology to do, even if it is somewhat different work than that advertised. In particular, and

[10] This characterization of Strong emergence departs in letter but not spirit from my preferred characterization, in ways related to the difference between a one-one and a one-many approach to metaphysical emergence, as discussed in Ch. 1, note 11 of my book; for present purposes nothing turns on the difference.

notwithstanding that debates over reduction and emergence have taken for granted conceptions of levels and their occupants making room, at a level, for structured as well as unstructured entities (and associated features)—such that, e.g., an atomistic physical level would contain not just atoms or pluralities of atoms, but also massively complex combinations of atoms standing in atomic relation—more work needs to be done as regards the details of how the domain of goings-on at a given level are generated. Boolean and classical mereological resources are also typically operative in generating 'lightweight' constructions of entities appropriately placed at a level, as I discuss in Ch. 1, Section 1.4.2, pertaining to the individuation of levels. Calosi and Giordani's system, and Calosi's attention to the difference between Reducible and Irreducible wholes, encode mereological resources which are both new and arguably 'lightweight.' These resources might well be added to the mix of those generating goings-on properly located at a level, and so be indirectly, if not directly, relevant to accommodating emergence, after all.

4. Replies to Emery

Emery's contribution raises a number of important questions stemming from an implemention of Weak emergence in terms of an elimination of degrees of freedom (DOF), of the sort I first offered in my 2010, and which plays a role in my book discussions of the emergence of complex systems (Ch. 5) and ordinary objects (Ch. 6). In this work, a DOF-based account is used to motivate the Weak emergence of certain special science goings-on from lower-level physical (e.g., quantum) goings-on. The overarching theme of Emery's questions concerns the extent to which attention to relations between DOF can be extended to address other cases of emergence—most interestingly, in my view, to cases of purported emergence within physics itself. A full treatment of Emery's unified set of questions deserves its own article; here I'll provide some initial response to what I see as her most pressing questions, and say a bit more about related questions in the footnotes.

Emery wonders, to start, whether a DOF-based implementation of Weak emergence might provide a fully general basis for Weak emergence—and if not, why not? To motivate my response to this question, it's worth recalling that my goal in the book is to consider whether, and ultimately to argue that, certain appearances of metaphysical emergence, drawn from both the special sciences and ordinary experience, can be taken at realistic face value. As such, I am looking to the sciences and to ordinary experience for input into which goings-on are, in those contexts and on the face of it, seemingly both dependent and autonomous in the ways characteristic of metaphysical emergence; and then my goal is to consider whether, and if so how, these appearances of metaphysical emergence can be taken at face value.

Now, a DOF-based implementation of Weak emergence reflects certain facts on the ground, including that certain special science entities are posited as having characteristic features encoded in associated special-scientific laws; that these entities are understood as composed by (systems of) lower-level entities which are also understood as having characteristic features encoded in associated (more) fundamental physical laws; and that the DOF needed to specify certain characteristic states of the former are eliminated as compared to the DOF needed to specify those same characteristic states of the latter. These facts, I argue, enter into a scientific

law-based motivation for thinking that some of the appearances of metaphysical emergence can be understood in terms of an elimination in DOF.[11]

Perhaps there are alternative ways of associating characteristic states of an entity with DOF which don't proceed by attention to scientific laws, in which case a DOF-based approach might be generalized to cover cases of Weak emergence involving such entities (or their features). But, two points. First, the availability of parameter-based accounts of characteristic states of, e.g., mountains, certain conscious mental states, or freely acting persons isn't obvious; in these cases other (e.g., functionalist or determinable-based) implementations of Weak emergence appear to be more naturally implemented.[12] Second, the conception of DOF as closely linked to certain laws plays an important role in my arguments for the conclusion that eliminations in DOF satisfy the conditions in the schema for Weak emergence, both in that the connection between DOF and laws is what blocks the reducibility of special science entities whose characterization involves eliminated DOF (since the lower-level laws require all the relevant DOF in order to operate), and in that insofar as laws express what an entity (system of entities) can or can't do, they also serve to encode what powers the entities have or don't have, in ways that suggest that entities with eliminated DOF as compared to the system of their composing/realizing entities will have fewer token powers than that system. Correspondingly, it's not clear that a DOF-based approach to Weak

[11] On this last, Emery also asks: what is necessary for a degree of freedom to count as eliminated? It can't be that the eliminated (e.g., quantum spin) degrees of freedom are *never* relevant to the behaviour of the entity at issue, for as I note in discussing the Weak emergence of ordinary objects from quantum goings-on, one can set up scenarios (e.g., a variation on Schrödinger's cat case) where quantum phenomena do impact the behaviour of the macro-entity. This question is related, in turn, to the question of which states, with associated DOF, are taken to be 'characteristic' of a given entity. Ultimately, I think that the answer depends on what makes for the sort of non-fundamental joint in nature of the sort plausibly encoded in special science laws. I don't have a general account of what makes for a non-fundamental joint, in part reflecting my view that there are many and diverse metaphysical dependence relations operative in cases of relative fundamentality (following my 2014 and elsewhere, to be given a broad defense in my forthcoming and under contract). But perhaps traction in the present case can be gained by attention to the usual view of special science laws as containing ceteris paribus laws or clauses, which allow for exceptions; and it might also be worth exploring (perhaps drawing on the degree-theoretic variation of the account of metaphysical indeterminacy advanced in Wilson 2013 and Calosi and Wilson 2018) whether Weak emergence comes in degrees, with non-fundamental joints being to some extent fuzzy or metaphysically indeterminate.

[12] A related line of thought applies to Emery's question of whether the DOF-based Weak emergence of ordinary objects might be gained, not by way of the elimination of quantum DOF (as I do in the book), but rather by way of the elimination of broadly statistical-mechanical DOF. Indeed, I appeal to thermodynamic features as having eliminated statistical-mechanical DOF in support of certain complex systems being Weakly emergent, reflecting the applicability of renormalization group methods to such systems when near critical points, which methods track the elimination of such DOF. I focus on the quantum case in the chapter on ordinary objects mainly because, again, there's a clear scientifically endorsed line of thought which (unlike applications of the renormalization group to gasses and other complex systems) targets ordinary objects—and not because a case for Weak emergence needs to involve an absolutely fundamental base; I agree that it doesn't.

emergence can be generalized to cases where no laws are operative without undercutting the motivations for the approach in the first place.[13]

To my mind, the most pressing of Emery's questions pertains to whether and how my schemas for metaphysical emergence might accommodate cases of such emergence within physics itself. As Chalmers (2021) observes:

> Discussion of "emergent spacetime" has exploded, driven largely by theories of quantum gravity—including versions of string theory, loop quantum gravity, and causal set theory—in which spacetime may not appear on the fundamental level. […] The key thesis is that spacetime exists at a nonfundamental level and is grounded in a fundamental level which is nonspatiotemporal. (164)

(See, e.g., Lam and Wüthrich 2018 and Huggett 2021.) Not just quantum gravity (QG), but general relativity (GR) itself (as presupposing relationism; see Rovelli 2007) and quantum mechanics (QM) (if the wavefunction/configuration space is taken as fundamental; see, e.g., Albert 2013, Ney 2021) have been taken to support spacetime as emergent in that spacetime is not fundamental, but is rather completely dependent on more fundamental nonspatiotemporal goings-on. Note that the supposition that emergent spacetime (or its three-dimensional occupants) are nonfundamental indicates that the type of emergence being posited here is of the Weak rather than Strong variety.

Such applications are in prima facie tension with my schemas for metaphysical emergence. One source of tension, observed by Emery, is that the notion of cotemporal material dependence often involves the composition of the entity having the emergent feature by lower-level dependence base entities; but in the cases at issue it is unclear how elements of the more fundamental physical ontology would 'compose' the emergent physical ontology (as Baron 2019 discusses; but see Baron and Bihan 2022 for an attempt to make sense of this). Now, my own view (as I register in my replies to Calosi, above) is that compositional relations aren't required for there to be emergence, but even so, one might be concerned that the dependence condition in the schemas is too restricted to make sense of cases of purported emergence within physics. Let's focus on the purported emergence of spacetime. Recall that the dependence condition encodes substance monism, whereby the only matter is physical matter, along with minimal nomological supervenience, whereby an emergent feature S requires and is at least nomologically necessitated by ('minimally supervenes on') cotemporal base-level goings-on P. As such, the dependence condition presupposes spacetime: S is

[13] Emery also wonders whether attention to DOF might enter into an implementation of Strong emergence, as involving a new DOF—and if not, why not? I didn't advance a DOF-based implementation of Strong emergence mainly because I didn't see clear case studies involving the posit of new DOF. As I discuss in Ch. 5 (182-5), so-called 'order parameters' are sometimes presented as involving new DOF (by, e.g., Morrison 2012 and Lamb 2015), but on closer examination no new DOF are really at issue: either the DOF are present at the micro-level, and what is new is their taking on certain values, or else the order parameters are not genuine DOF, but rather 'phenomenonological descriptions' of a system's order. That said, if there were cases where an apparently new DOF could not be given a reductive or other deflationist treatment, and given that the new DOF was associated with behaviours, law-governed or not, then a DOF-based implementation of Strong emergence might well make sense—though in such a case it's not clear that we would be adding anything new beyond the existing claim that a Strongly emergent feature has a fundamentally novel power.

cotemporal with P, and entities possessing these features will typically (per substance monism) share matter, hence spatially overlap.

The autonomy conditions in my schemas also presuppose spacetime: though a power may never be exercised, in any case powers are had by, and causal relations hold between, phenomena which are spatially located; moreover, accounts of causation take this to be either diachronic or synchronic, and so presuppose the notion of time.[14]

My conceptions of metaphysical emergence are not unusual in taking spatiotemporal notions for granted; effectively all standard conceptions do so. Those exploring the status of spacetime as emergent typically recognize that there is a prima facie difficulty in taking ST to be emergent by lights of standard accounts, and in response weaken the notion of emergence by removing references to space or time. There are a couple of different strategies on offer here, but in my view it is not clear that these attempts succeed—effectively, because satisfaction of the weakened conditions is compatible with either reduction or with Strong emergence, contrary to the intended characterization in these suggestions of spacetime as Weakly emergent from more fundamental nonspatiotemporal ontology.

One sort of strategy involves characterizing the dependence and autonomy conditions in ways eliding reference to spatiotemporal notions, as in Crowther's (2022) characterizations in terms of

1. dependence (cashed in terms of asymmetric supervenience correlations)
2. novelty (cashed in terms of qualitative difference)
3. autonomy (cashed in terms of multiple realizability or determination)

Crowther distinguishes 'hierarchical' emergence (a non-ST form of cotemporal emergence) where the base is somehow present, and 'flat' emergence (a non-ST form of diachronic emergence) where ST results from a non-causal 'interaction'.[15] And she argues that on certain accounts of quantum gravity, spacetime satisfies the dependence and autonomy conditions vis-à-vis the specified non-spatiotemporal basis, in hierarchical or flat fashion (and maybe both).

But granting satisfaction of these conditions in some or other versions quantum gravity, the associated weakened conception of emergence is too weak to establish that spacetime is less fundamental than what it depends on:

> A's asymmetrically supervening on B doesn't entail that A is less fundamental than B.[16]
>
> A's being qualitatively different with respect to B doesn't entail that A is less fundamental than B.
>
> A's being multiply realized/determined by B, C, and D doesn't entail that A occupies a less fundamental level than B, C, and D; for if A is identical to the disjunction of B, C, and D (as reductionists typically maintain), A will be as fundamental as the disjunction. That is reduction, not emergence.

[14] "The lack of a metric structure […] seems to result in the loss of causation since, on the face of it, causation requires (at least) time to exist. […] Causation is usually thought to be a relation between events, which are individuated by their spatiotemporal locations" (Baron 2019: 2208). That said, some recent conceptions of causation do not build in the notion of time; see Baron and Miller 2014 and Tallant 2019.

[15] Here the model is something like the occurrence of the big bang.

[16] For example, determinables asymmetrically supervene on determinates; but many think quantum determinables are prior to their determinate values.

> A's being multiply realized/determined by B, C, and D is also compatible with the base phenomena serving as diverse preconditions for something fundamentally novel. That is Strong, not Weak, emergence.

As such, as they stand Crowther's conditions on emergence are too weak to rule out either (identity-based) reduction or Strong emergence. Moreover, on the face of it this weakness reflects the elision of spatiotemporal notions from these conditions. The best shot for establishing genuine autonomy of a Weak emergent variety proceeds by resisting reductionist and Strong emergentist readings by attention to causal considerations, and more specifically via satisfaction of the proper subset of powers condition, which blocks Strong emergence, since such emergence requires a novel power, and blocks reductionism, since disjunctive features are instanced by instancing a disjunct, and each disjunct has more token powers than are had by the Weakly emergent feature. As above, such causal notions appear to presuppose spacetime, and this is true as well on a DOF-based implementation of Weak emergence. That said, in other work Crowther (2018) suggests that a DOF-based implementation of Weak emergence can make sense of the emergence of spacetime from at least some nonspatiotemporal fundamental ontologies:

> Wilson's (2010) weak ontological emergence, where an emergent theory may be characterised by the elimination of degrees of freedom from the underlying theory […] is certainly applicable if spacetime emerges as illustrated by the condensed matter approaches to QG, and it applies to GFT, and any other approaches where spatiotemporal degrees of freedom emerge as collective, low-energy variables, analogous to those of thermodynamics. It also may apply in the context of LQG, where degrees of freedom possessed by the spin foams are eliminated in the approximation and limiting procedures designed to resolve and/or wash-out their discrete nature and quantum properties in the recovery of spacetime. (84)

These are intriguing suggestions. If Crowther is correct, and in a way I hope she is, then I would need to back off, at least for the case of spacetime, from the claim in my book that satisfaction of the conditions in the schemas is 'core and crucial' to metaphysical emergence of the sort connecting special science and fundamental physical goings-on. I'd need to say something more general.[17] Though my arguments that eliminations in DOF suffice to block reductionism and Strong emergence presuppose that DOF are associated with broadly causal laws, perhaps the same line can be implemented using a non-causal notion of information. This is something I'm working on. At present it's not entirely clear to me that there is a workable conception of Weak emergence—one which ensures dependence with autonomy—that abstracts away from causal or other spatiotemporal considerations.

A second strategy aimed at accommodating the emergence of spacetime involves appealing to a specific relation as holding between spatiotemporal and non-spatiotemporal ontology, suitable for seeing the former as dependent yet autonomous from the latter. Here the most popular suggestion appeals to something like functional realization:

> On a functionalist picture, whether an entity (a structure, object or property—from now on I will just say "structure") counts as spatiotemporal is determined by its

[17] Or disjunctive—but that would be unsystematic.

> functional role. The functional role of a physical structure is its role in the physical laws, which often boils down to its implications about the motion of material objects. (Baker 2020: 278)

This suggestion is subject to the sort of considerations I discuss in my book when discussing functional realization in special-scientific contexts. To start, we must distinguish between 'realizer' functionalism, on which functionally implemented goings-on are identified with the realizer of the role, and 'role' functionalism, on which functionally implemented goings-on are identified with the role itself, usually understood as a kind of higher-order property. Realizer functionalism is compatible with (indeed, is a form of) identity-based reductionism, and so is unsuited for purposes of vindicating the metaphysical emergence of spacetime from nonspatiotemporal ontology. Role functionalism potentially does better; and here (following the literature in metaphysics of mind/science), what's needed is some reason to think that there exists such a second-order feature. And the usual means of doing this is by appeal to the multiple realizability of spacetime. But as I'm at pains to highlight in my book, a mere appeal to multiple realizability does not suffice to establish the irreducibility of the multiply realized feature. In particular, work must be done to rule out a disjunctive treatment of the multiple realizability at issue. And again, the main strategy for doing this (mine) appeals to causal considerations, so won't work here—though it may be that looking to eliminations of DOF is the best bet here.

But suppose it turns out that no implementation of a (nonspatiotemporal) variation of my schemas for metaphysical emergence can make sense of the purported emergence of spacetime (or its occupants). In that case, I'll here register that there are alternative, and to my mind more natural, ways of thinking about some of the relations between nonspatiotemporal and spatiotemporal goings-on than in terms of metaphysical emergence. In particular, we have in hand certain metaphysical conceptions of how concrete goings-on are related to comparatively abstract goings-on, including ones on which abstract universals (not in space and time) come to be concretely instantiated, and ones on which among the space of abstract possibilities (not in space and time), just one comes to be actualized. This last seems especially relevant to the present case; for if (following Allori) the wavefunction represents possible ways the world or objects in the world can be, then configuration space is properly seen as a modal space, with concrete goings-on being best understood as instantiations or actualizations of these possibilities. These relations—instantiation, actualization—deserve further investigation and attention. For present purposes, what is important is that there is no clear sense in which the instantiation of a universal, or the actualization of a possibility, is any less fundamental than the universal/possibility. So why think that the relation between configuration space and ordinary spacetime and its occupants entails that the latter is less fundamental than the former? Either way, the relation isn't one of metaphysical emergence per se—in which case the inability of an account of metaphysical emergence to apply to these cases doesn't pose a problem for the account. But again, as with other of the questions Emery raises, there is more work to be done in arriving at a considered answer.

5. Replies to Gozzano

Gozzano's comments address the interesting question of whether the common supposition that Weakly emergent mental features are multiply realizable—or as

he puts it, are 'realization indifferent'—is compatible with the plausible supposition that mental features are 'systematic', in entering into patterns of dependencies. Gozzano expresses the potential threat to mental features' being Weakly emergent in the form of an argument:

> (i) Mental features are systematic;
> (ii) (In many cases) Emergence entails realization indifference;
> (iii) Systematicity entails that realization indifference cannot hold;
> (iv) Therefore, (in many cases) mental features can't be emergent. (271)

(Gozzano puts aside Strong emergence, as implausible; hence here and elsewhere his references to 'emergence' are more specifically to Weak emergence.) Each premise in this argument, Gozzano maintains, can be defended; and the conclusion therefore follows.

The focus of my response in what follows is on premise (iii), but let me start by saying a bit more about (i) and (ii).

First, in re the claim that mental features are systematic. Gozzano doesn't offer a definition of 'systematicity', but does offer a number of illustrations, including cases where increases in the intensity of a perceptual stimulus are (e.g., logarithmically) systematically associated with the intensity of the phenomenal state, and cases where changes in the intensity of a phenomenal state (e.g., pain) are systematically associated with an increase in some other phenomenal state (e.g., anxiety). Though I'm not sure about the status of these particular examples, I think that what Gozzano has in mind here is that there might be relations—better, to avoid ambiguity, 'mappings'—between (to speak loosely) families of mental feature types whereby members of one family are systematically related with members of the other family. I'm happy to grant that various special science laws, including those of psychology and neuropharmacology, will at least sometimes encode these sorts of systematic mappings between (families of) mental features.

Second, in re the claim that many cases of emergence entail realization indifference, three observations. To start, I'd prefer 'involve' over 'entail', since whether a given higher-level feature is multiply realizable is an empirical, not logical, matter. Next, Gozzano's discussion involves a characterization of 'realization indifference' as building in the possibility of 'wildly different' realizers; this goes beyond the usual appeals to multiple realizability as motivating Weak emergence, which appeals often involve realizers being only 'mildly' different, as when, e.g., my belief that Paris is beautiful is realized by different neurological states, or the shape of a flock of birds is realized by different configurations of its constituent birds, will do. As such, in what follows I will usually revert to the usual terminology of multiple realizability, but will revisit whether the possibility of 'wildly different' realizers makes any difference down the line. Finally, as Gozzano notes, I don't take multiple realizability to be either necessary or sufficient for weak emergence: not necessary, since there are cases to be made that some singly realized features satisfy the proper subset condition on powers; and not sufficient, since reductionists have strategies for accommodating multiple realizability in disjunctive or other terms, which must be blocked before multiple realizability can be assumed to involve emergence. All this said, Gozzano is right that many cases of Weak emergence, of mental features in particular, are initially and primarily motivated by multiple realizability; so it is definitely worth considering

whether these suppositions are in tension with the also-plausible assumption that mental states enter into systematic mappings.

I now turn to the key premise (ii) in Gozzano's main argument—namely, the claim that 'Systematicity runs against realization indifference'.

The underlying motivation for Gozzano's endorsement of this claim appears to be a supposition that if special science properties enter into systematic mappings, then the lower-level properties upon which the special science properties cotemporally materially depend must also stand in systematic mappings. As he puts it,

> if we consider the causal relations in which [systematic special science feature] *S* is involved [as encoded in] high-level laws of the sort discussed by special sciences, we may require a sort of systematic counterpart of supervenience: there cannot be systematic variations at a high level without systematic variations at a low level. (272)

In this sense, Gozzano supposes, the existence of a systematic mapping between (families of) higher-level features places constraints on the realizers of these features—constraints which, he maintains, are not in place in cases of multiple realizability.

Gozzano offers a specific subargument in support of this claim and the associated premise in his main argument. In the interest of efficiency I will focus my critical attention primarily on a key premise (2) in that subargument, according to which (and consonant with the previous line of thought), if a property *S* is systematic,

> (2) The *P*s on which *S* cotemporally materially depends (CMD), should follow the same pattern of systematicity shown by *S*. (274)

Now, it is unclear why we should accept this. As Gozzano observes:

> One may wonder why the emergentist should accept [this] premise […]. The emergentist can stress that each "level of reality" […] is characterized by its laws […] and on which *S* cotemporally materially depends. So, what consequences would bear [on] having different systematic relations, if any at all? (274)

The complaint here seems to me to be apropos, as far as it goes. Even granting that systematic mappings between (families of) higher-level features requires systematic mappings between (families of) lower-level features, why would these mappings have to 'follow the same pattern'? Indeed, it's not clear that higher-level systematicity mappings require lower-level systematicity mappings. All that ultimately seems required to accommodate systematic mappings involving realized features is that their lower-level realizers enter into laws compatible with those higher-level systematic mappings. Maybe those lower-level features and laws will also fall into 'systematicity patterns', but at the end of the day all that's required is that any given realizer of any given higher-level feature *S* provide a suitable basis for *S*'s having the powers it needs to have to conform to whatever systematicity mappings are in place.

So, Gozzano's premise is better expressed as requiring not that realizers enter into the 'same pattern of systematicity' as *S*, but just that (at most) whatever laws are in place as regards *S*'s realizer on a given occasion serve as an appropriate

basis for accommodating the systematic mappings into which S enters. To assess whether systematicity runs against multiple realizability, then, the question is whether there are reasons to think that a feature's being multiply realizable somehow poses a problem for its realizers' accommodating the systematic mappings into which S enters.

I answer in the negative; I don't see any problem here. Since at issue are cases where multiple realizability ends up motivating Weak emergence, let me put the point in my favoured terms. To fix ideas, suppose that mental features M_1, M_2, and M_3 are systematically causally connected to mental features M'_1, M'_2, M'_3; suppose also that each of these six types of mental features is multiply realizable; and suppose that (after undercutting reductionist strategies) this multiple realizability is taken to support these features' satisfying the conditions on Weak emergence vis-à-vis whatever features realize them on a given occasion. Here the systematic mapping (like Gozzano's illustrative cases) causally connects certain mental features with certain others; hence to accommodate this mapping just requires that any realizer of M_1 has among its powers the power to cause M'_1, any realizer of M_2 has among its powers the power to cause M'_2, and so on. But on the operative understanding of realization, this follows automatically, since any token power of a realized (Weakly emergent) feature on a given occasion is identical to a token power of the feature that realizes it on that occasion. So the treatment of M_1 as both multiply realized and Weakly emergent is compatible with M_1's entering into the systematicity mapping; and similarly for M_2 and M_3. So systematicity is here accommodated, notwithstanding the multiple realizabilities of the features at issue.

Note also that we were able, in this narrative, to remain neutral on whether the realizers of the mental features themselves enter into a systematicity mapping, whether similar to or different from those into which the mental features enter. Whether this is so will depend on further details about the powers and power profiles of the realizers. This bears on premise (5) of Gozzano's subargument according to which "If [the realizers] have different projectability patterns and support different counterfactuals, they do not establish the same systematic relations" (274). To be sure, the realizers can be expected to enter into different projectability patterns and support different counterfactuals (it is precisely this difference that provides a basis for thinking that Weak emergents are distinctively efficacious, in spite of not having any new powers), and let's even grant that the realizers themselves don't enter into systematic mappings at all, much less 'the same' ones into which mental states enter. None of those further details matter for whether multiple realizers can accommodate higher-level systematic mappings, as the previous case illustrates. All that matters is that the realizers have the requisite powers—as they will do, on my account of Weak emergence.

This seems to me to be a coherent narrative, indicating that there is no in-principle problem with there being systematic, multiply realizable mental (or other) features.

It remains, however, to consider two strategies for defending Gozzano's claim to the contrary. The first reflects Gozzano's characterization of multiple realizability as realization 'indifference', such that the diverse realizers at issue may be 'wildly' different—so different that they might share nothing in common:

> Let's consider pain: supposedly, in humans, it is realized by C-fiber firing, but it could be differently realized in other sentient beings and the realizers form an open set. So, we may take the property of being in pain as one that at a very high level

> can be shared by different entities, from human beings to other mammals, to other animals up to potentially extra-terrestrial individuals. At a finer level of detail, being in pain is multiply realized by structures that may have nothing in common. (270)

Supposing it were the case that diverse realizers of a single feature might have 'nothing in common'—in the case of systematically related mental features, in particular—then I can see how Gozzano might conclude that systematicity runs against realization indifference. My response here is simply that I reject the supposition that realizers might 'have nothing in common', since that supposition leaves it unclear how or in what sense one feature might realize another. If, as I argue is the case for the broadly scientific (including mental) features that are the target of my book, the feature whose realization is at issue has a distinctive power profile, then at a minimum any realizer of a feature *S* has to have, among its powers, the powers of *S*. (And as I also argue, a wide range of accounts of realization, including functional realization and the determinable-determinate relation, agree.) On such an understanding of realization, effectively encoded in the schema for Weak emergence, this much will be 'in common' among multiple realizers of a feature, and as per the case above, that much seems sufficient unto the task of accommodating systematicity.

The second strategy pushes in a different direction, and is suggested by Gozzano's discussion of what powers should be taken to be in the power profile of a given feature:

> According to the subset strategy a property is individuated by the set of its causal powers had by all its instances […] But the causal powers defining the set do have causal relations to other powers. Say, a rubber band is elastic and green. Elasticity is shared among all elastic entities no matter their color. But elasticity determines fragility in cold conditions. Should we consider this as a condition on other elastic entities? […] Should the elasticity also involve a specific ratio between, say, thickness and length of stretchability? If so, then it could be the case that only a specific realizer fits the bill. But if this is the case, then it seems Kim was right after all: each disjunct has its own merits and the high level is just a measure of our ignorance. (275)

Here one can see Gozzano as maintaining that closer examination of the powers associated with a given property indicates that powers are much more finely individuated than is usually recognized, to the extent that the claim that features, including those entering into systematic mappings (which impose yet further constraints on powers) are not appropriately seen as multiply realizable. My response starts by observing that, although this is often mainly left tacit for simplicity, talk of 'powers' in these contexts is intended as talk of 'conditional powers', such that powers are individuated not just by their effects, but also by the intrinsic and extrinsic conditions required for the powers to be manifested or exercised. Hence any given property will be associated with massively many conditional powers—not just 'the power to stretch without breaking', but 'the power to stretch without breaking if instantiated in warm conditions', and so on. All these conditional powers are had by any instance of a feature, even if the conditions of manifestation of the power do nor or even cannot obtain (as when a plastic knife has the property of being knife-shaped, which includes among its powers the power to cut wood if made of steel). This understanding strikes me as unifying and systematic,

and in line with the connection between (in particular) scientific taxonomy and laws, so I am inclined to stick with it, rather than adopting such a fine-grained conception of powers that hardly any features turn out to be multiply realizable.

To return to Gozzano's primary argument: since I can reasonably deny that 'Systematicity runs against realization indifference', I can resist Gozzano's conclusion that considerations of multiple realizability don't support the Weak emergence of mental features—especially those entering into systematic mappings.

That said, I want to close by registering that Gozzano has called something important to attention—namely, that broadly holistic considerations may turn out to be relevant to discussions of metaphysical emergence. Discussions of metaphysical emergence have tended to focus on individual cases—this mental feature, that thermodynamic feature, and so on. But how do systematicity mappings and other more global considerations bear on this topic? For example, in the case above, might M_1 and M_2 be Weakly emergent and M_3 Strongly emergent, or is there some reason to think that systematically related features should, or even must, have the same status? This is a new question, and deserves further attention.

6. Replies to Onnis

In *Metaphysical Emergence*, I motivate my powers-based schemas for Weak and Strong metaphysical emergence by attention to Kim's problem of higher-level causation, which I present as "the most pressing challenge to taking the appearances of emergent structure as genuine" (39). Onnis's contribution is aimed not at directly problematizing the schemas themselves, but at calling into question their underlying motivation in Kim's problem of higher-level causation. She aims to argue that Kim's argument proceeds against certain metaphysical presuppositions—each associated with 'Alexander's Dictum', according to which to be real is to have causal powers—which, if rejected or differently interpreted, would render the argument less of a challenge so far as accommodating emergence is concerned. As she summarizes:

> [T]here are three issues that need to be addressed. The first one concerns the Dictum itself: one may want to reject it and assume other criteria about existence. The second one is about the power-based interpretation of the Dictum: one may want to accept the latter, while considering its power-based interpretation as too strict. The third one is about the metaphysical underdetermination of the powers involved in the power-based interpretation: one may want to accept the Dictum and its power-based interpretation, while requiring a differentiation between microscopic physical powers and macroscopic emergent powers. (296)

Since the problematic presuppositions at issue concern powers, one can see Onnis here as pushing back not just on the stated motivations for my schemas, but more pressingly on my claim that the powers-based schemas are 'core and crucial' to metaphysically accommodating the appearances of emergence.

The considerations that Onnis raises are well worth attention. Even so, as I will now argue, at the end of day the metaphysical presuppositions she identifies as underpinning Kim's problematic are not required for this problematic to put pressure on the viability of metaphysical emergence; hence the motivation for my powers-based schemas as indeed 'core and crucial' to accommodating such emergence remains.

6.1 Alexander's Dictum

As discussed in the *Précis*, I set out Kim's overdetermination problem as involving six premises, four of which (Reality, Distinctness, Efficacy, and Dependence) encode certain assumptions about the seeming higher-level features at issue, and two of which (Physical Causal Closure and Non-overdetermination) encode certain assumptions about causation. The basic concern is that any purported effect of a (real, distinct, dependent) higher-level feature is (per Closure) already brought about by the lower-level physical goings-on upon which the higher-level feature depends, and so is (contra Non-overdetermination) overdetermined. As I observe, standard responses to Kim's argument are associated with certain views, denying some or other premise. Of these views, only those denying Physical Causal Closure (i.e., British emergentism) or Non-overdetermination (i.e., non-reductive physicalism) accommodate metaphysical emergence, understood as coupling cotemporal material dependence with ontological and causal autonomy (distinctness and distinctive efficacy); and the strategies encoded in these two views motivate my schemas for emergence, whereby a higher-level feature has a fundamentally novel power as compared to its dependence base feature on any given occasion (Strong emergence), or a higher-level feature has a proper subset of the token powers of its dependence base feature on any given occasion (Weak emergence).

Now, Onnis maintains that Kim's overdetermination argument presupposes Alexander's Dictum (after British emergentist Samuel Alexander), commonly spun (e.g., by Kim 2006: 557) as the thesis that 'to be is to have causal powers'. To start, Onnis observes, Kim takes Alexander's Dictum to motivate the Efficacy premise in his argument (perhaps given the Reality premise in his argument), insofar as he registers that "to be a mental realist […] mental properties must be causal properties" (1998,: 43). Moreover, in his (2006), Kim goes further, saying "Properties that are lacking in causal powers—that is, whose possession by an object makes no difference to the causal potential of the object—would be of no interest to anyone" (557), again connecting this thesis to Alexander. Onnis goes on to claim that Kim's problem requires and gets traction only under the assumption of Alexander's Dictum:

> If the principle is rejected, entities can have a legitimate existence even without exerting causal efficacy. If the nonreductive physicalist has to give up her nonreductionism, therefore, it is because of Alexander's Dictum. (292–93)

I respond that it isn't clear either that Kim accepts Alexander's Dictum, or that Kim's problem gets traction only if one assumes this Dictum. As regards Kim's own proclivities, it is worth noting that his expressions of claims in the ballpark of Alexander's Dictum (as in his 2006, above) are uniformly offered in a context within which he is presenting the emergentist's point of view, as opposed to his own. In any case, Alexander's Dictum is very broad; it aims to provide a general necessary condition on the existence of goings-on of any ontological category whatsoever. As such, one might reject the Dictum in full generality—perhaps because one believes that platonic universals or numbers exist, but don't have causal powers—yet still maintain that for scientific or concrete entities and features, to be is to have causal powers. Indeed, Kim's focus in his discussion of overdetermination is squarely on broadly scientific features, so it isn't obvious that he intends

to advance anything as strong as Alexander's Dictum, understood as a general criterion of existence.

That said, Onnis is correct that Kim's problematic takes as a premise that mental (more generally: special scientific) features have powers, as per Efficacy. But we don't need Alexander's Dictum to motivate this premise. Independent of that Dictum, the efficacy of special science features is motivated by their entering into special science laws which standardly express causal regularities (chemical reactions, geological forces, biological processes, predator-prey relationships, neurological and psychological interactions, and so on). And we moreover have direct experience of the seeming efficacy of the qualitative mental features that are Kim's primary focus, as is reflected in nomological truisms such as that (ceteris paribus) being in pain causes avoidance behaviour, being hungry causes one to seek out food, and so on.

These independent motivations for taking the higher-level features at issue in Kim's problematic to be efficacious would remain even if one rejected Alexander's Dictum, either in full or in part, perhaps on grounds (as Onnis suggests) that certain motivations for thinking that some goings-on exist don't explicitly require the efficacy of said existents.[18] It would remain that there are theoretical (law-based) and experiential reasons for thinking that mental and other special-scientific features are efficacious; and given the other premises in Kim's argument, his challenge for there being emergent higher-level goings-on would unfurl accordingly. To be sure, the epiphenomenalist responds to Kim's problematic by denying Efficacy; but to offer an epiphenomenalist response to Kim's problematic is not to say that there was never a problematic there in the first place. On the contrary, in the dialectical course of events the burden is on the epiphenomenalist to explain away the science-based and experience-based motivations for Efficacy—a burden not easily discharged, which may account for the relative paucity of epiphenomenalists.

6.2 A Heavyweight Notion of Powers?

I next turn to Onnis's claim that, even granting Alexander's Dictum (at least as applied to mental and other scientific features), Kim's interpretation of the Dictum presupposes a conception of efficacy as involving powers that are real in some metaphysically heavyweight sense. As Onnis interestingly argues, such a conception appears to be at odds with Alexander's own comparatively lightweight correlational conception of efficacy. She moreover suggests that a heavyweight conception of powers "seems to already carry anti-emergentist implications", insofar as such powers are a ready target of reductionist strategies. For example, Onnis observes that on one implementation of Taylor's (2015) 'collapse' objection to the viability of Strong emergence, any purportedly fundamentally novel powers at the

[18] By way of such alternative motivations, Onnis considers being introspectively accessible (as I suggest provides defeasible motivation for our taking libertarian free choice to exist) or being indispensible to our best science. Introspection of free will seems to me to satisfy Alexander's Dictum twice over, insofar as a free choice causes both the awareness of the choice and the outcome of the choice. Indispensibility considerations look better by way of a genuine alternative motivation for existence—perhaps causally inert mathematical entities are required for our best theories. In any case, the availability of such alternative motivations doesn't undercut the specifically causal considerations which motivate mental and other special-scientific goings-on.

higher level can be traced to dispositional properties of base-level constituents. Onnis suggests that less committal conceptions of the efficacy at issue "seem to make the problem of higher-level causation less challenging".

It is true that Kim frames his problematic in terms of powers, as in his Causal Inheritance principle and elsewhere. So far as I can tell, however, all that Kim has in mind in his talk of 'causal powers' associated with a given property is that the having of the property 'makes a difference' to the causal potential of an object—that is, to what the feature (or an object having the feature) can cause when in certain circumstances. Such an understanding is in line with the metaphysically neutral understanding of powers operative in *Metaphysical Emergence*, according to which talk of powers is talk of the contribution that the having a property can make, when instanced in appropriate circumstances, to the production of a given effect. This neutral understanding does not require that powers be understood as dispositions or in any other heavyweight terms; as I argue (33), even a contingentist categoricalist Humean could accept powers in the sense operative in the schemas.

In any case, suppose that the operative notion of efficacy/causation and associated use of 'power' is given a weak—say, Humean—reading in Kim's problematic. Would Kim's argument then pose less of a threat to accommodating the appearances of higher-level reality, as involving emergent special science features? One motivation for a positive answer might proceed as follows. To start, consider the sort of scenarios that are not supposed to be good models for making sense of higher-level causation: namely, firing squad or double-rock-throw cases. Why think that it would be problematic if mental or other special science causation were overdetermined like this? The concern seems to reflect a kind of 'oomphy' understanding of efficacy, where different causes directed at the same effect would, like different substances trying to occupy the same space, get in each others' way. And perhaps such an 'oomphy' understanding is more naturally associated with a heavyweight notion of powers, as real dispositions or the like.

But even supposing a more metaphysically substantial notion of efficacy or power provides one route to finding causal overdetermination problematic, it isn't the only way. Another route simply lies in observing that, whatever the right account of causation, and whatever (in particular) is going on in firing squad and double-rock-throw cases, it remains that mental causation is *not that kind of case*—the relation between the mental goings-on and their physical dependence base is just different from those sorts of overdetermination cases. And yet certain of the premises in Kim's argument suggest that higher-level features would overdetermine the effects of their lower-level bases. That's really all that the 'Non-overdetermination' premise is registering; and Humeans as well as non-Humeans can and typically do agree that this is enough to get the problematic going.

Moreover, just because one accepts a Humean or other lightweight understanding of causation and associated talk of 'powers', or prefers to dispense with talk of powers altogether (even as shorthand for saying what can cause what), it isn't clear that the problem of higher-level causation thereby becomes less challenging. As I observe in Wilson 2002, if causal power is understood just as a matter of nomological sufficiency (in the circumstances), then insofar as base-level properties are nomologically sufficient for higher-level properties, and nomological sufficiency is transitive, then any power purportedly had by the higher-level

property will also be had by the base property.[19] Hence a version of the Collapse objection against Strong emergence attaches even to a lightweight conception of efficacy/powers.[20] And as I also observe in Wilson 2002, if causal power is understood just as a matter of nomological necessity (in the circumstances), then in any case where the higher-level property is multiply realizable, then the physical base-level property will be ruled out as efficacious.[21] In that case it would appear that Physical Causal Closure is violated, and Kim's problematic again comes into play, illustrating a prima facie challenge in reconciling higher-level causation with a broadly physicalist world-view.

So the force of Kim's problematic overdetermination argument does not hinge on commitment to a heavyweight conception of efficacy or powers. Luckily, or so I argue in my book, physicalists and non-physicalists alike have the resources, either in general or via appropriate implementations of the schemas for Weak and Strong emergence, to respond to the full range of ways in which Kim's challenge may be brought to bear.

6.3 Microscopic vs. Macroscopic Emergent Powers

Finally, I turn to Onnis's claim that taking there to be a "difference in kind" between higher-level and lower-level powers "might be able to weaken the problem of high-level causation":

> By examining the nature of causal powers, for instance, it might be discovered that higher-level powers cannot really collapse, while lower-level ones cannot really emerge. Emergent and non-emergent causal powers, in other words, might simply be non-interchangeable powers of a different kind. (300)

Onnis goes on to offer a preliminary characterization of the difference between 'emergent' and 'non-emergent' powers. The latter, she suggests, are associated with properties of micro-objects (e.g., the mass of an electron), and are commonly thought to be "fundamental, essential, intrinsic, intrinsically active, and productive". The former are associated with properties of macro-objects (e.g., the hardness of a diamond), and "are often conceived as nonfundamental, extrinsic, context-sensitive, and constraining", as on Gillett's (2016) understanding of 'machresis' as a form of non-productive 'role-shaping' determination. Onnis speculates that "the most striking difference between micropowers and emergent powers would therefore be the intrinsic activity and productivity of the former and the

[19] As I there illustrated: "[S]uppose one of my brain properties necessitates one of my mental properties, and the mental property bestows some causal power on me. [If] causal power bestowal is just a matter of nomological sufficiency, my brain property will, in virtue of necessitating the mental property, also bestow this causal power on me" (Wilson 2002: 64).

[20] I respond to this and other versions of the Collapse objection in my book (drawing on Wilson 2002 and Baysan and Wilson 2017), but the present point is just that the threat of Collapse does not hinge on a heavyweight conception of efficacy/powers.

[21] As I there illustrated: "The general idea is this: suppose either of two of my brain properties is sufficient for one of my mental properties, and the mental property bestows some causal power on me. Since we're assuming that causal power bestowal is a matter of nomological necessity, as well as sufficiency, and since neither brain property is necessary for the effect in question, neither brain property will bestow this causal power on me" (Wilson 2002: 65).

extrinsic non-productive constraining capacities of the latter" (300). And re the Collapse concern, she suggests that

> differentiating between micropowers and macropowers might make this collapse more difficult. For instance, let's suppose that the macroscopic causal powers exerted by a biological complex system require a biological complex bearer. In that case, a nonbiological system or a biological isolated component could not instantiate those macropowers, which would therefore become non-collapsible. (300)

Onnis notes that these suggestions are preliminary, but even so let me say why I'm not inclined to take on board any such distinction in kinds of powers. To start, I don't speak of 'emergent powers' (or non-emergent powers); it is features, or perhaps entities having the features, which are emergent (or not) on my view. And as above, the conception of 'power' operative in my book encodes just that (talk of) powers associated with a given feature is (talk of) what contributions the having of the feature may make to the production of certain effects, when in certain circumstances. Such a neutral characterization makes sense, so far as I can tell, whatever sort of feature or entity is at issue. Nor would I be inclined to endorse a conception on which emergent and non-emergent features (or associated powers) differ in fundamentality status, both because Strongly emergent features (powers) are just as fundamental as whatever fundamental physical features (powers) there might be, and because the physical features (powers) serving as a co-temporal dependence base for higher-level features (powers) will themselves typically be features of highly complex micro-configurations, and so not themselves be fundamental. I would also resist any general characterization of emergent features (powers) as 'constraining', not just because cases of Strong emergence needn't involve constraints, but also because cases of Weak emergence needn't do so (as on a determinable-based implementation); and even when Weak emergence does involve constraints, it is lower-level goings-on, not higher-level powers, which impose the constraints (as on the degrees-of-freedom-based implementation discussed in §5.2.4 of my book).

That said, I agree with Onnis that further investigations into the nature of powers might open the door to new strategies for responding to at least some concerns about emergent features. Indeed, Onnis's suggested response to the Collapse objection is quite similar to the 'new bearers' strategy which I discuss in Ch. 4 (135), which appeals to Baysan's (2016) view that features have their powers derivatively on the powers of their bearers. But note that whether one wants to go this route to avoid Collapse will depend on whether one is inclined to accept Baysan's view (which as it happens, I'm not). Moreover, the question will remain of whether the macrofeatures (powers) at issue in a given case are or are not in line with physicalism—which brings us back to the terrain of Kim's problematic.

To sum up: while it's worth asking whether Kim's problematic is generated by Alexander's Dictum or related controversial assumptions, my general answer is 'no, it isn't so generated'; and similarly for the Collapse concern for Strong emergence. Rather, these problematics are surprisingly robust across heavyweight and lightweight conceptions of efficacy and powers. As such, for those aiming to realistically accommodate the appearances of metaphysical emergence, the powers-based responses encoded in the schemas for Weak and Strong emergence remain the only game in town.

7. Replies to McLaughlin

In his contribution, McLaughlin raises several important questions about or concerns for my views. My responses here will focus on the following: first, whether my 'no fundamental mentality' account of the physical needs to embrace further constraints; second, whether satisfaction of the conditions in the schema for Weak emergence is either necessary or sufficient for physical acceptability; and third, whether Strong emergence, understood as involving fundamental powers or associated interactions which come into play only at certain levels of compositional complexity, is compatible with quantum field theory.

I start with a quick clarification. McLaughlin describes my account of the physical as one according to which the physical "[…] is whatever would be posited by the completed physics in fact true of our world, with the following caveat: A mental feature is not to be counted as a physical feature even if that physics would posit it" (280); and he describes the associated constraint on physicalism as one according to which "any doctrine deserving of the name "physicalism" should be incompatible with the physics in fact true of our world having to posit mental phenomena" (280). If by a 'posit' of physics we just have in mind the (most) fundamental entities or features treated by that theory, then these descriptions coincide with my account of the physical and the associated constraint on physicalism, respectively. But since physics also in some sense posits non-fundamenta (e.g., protons and other particles composed of quarks) and more generally treats certain non-fundamental complexes (e.g., pluralities or relational aggregates), it's worth being clear that what I rule out as 'physical' are any goings-on that are (as I put it) 'fundamentally mental', in being both (a) fundamental and (b) individually such as to have or bestow mentality, of the sort, e.g., that panpsychists suppose exist—hence the 'no fundamental mentality' (NFM) constraint. The NFM account is compatible, e.g., with physics treating non-fundamental physical states (consisting of some massively complex combination of fundamental physical goings-on) that are either identical with (as on a reductive physicalist view) or which realize (as on a non-reductive physicalist view) mental features.

Now, in re my NFM account of the physical, McLaughlin considers whether I would accept further constraints on the physical—e.g., a 'no fundamental chemical' and 'no fundamental biological' constraints—and speculates that I would do so:

> I think [Wilson] would […] accept such additional constraints. It is clear, for instance, that if the physics in fact true of our world would have to posit entelechies or a fundamental vital force, she would take physicalism to be false. (280)

McLaughlin doesn't present the potential need to introduce further constraints as an objection, but other things being equal, I would prefer not to introduce such further constraints, since it seems to me that doing so would be unsystematic. As I earlier put it:

> One might wonder whether imposing the NFM constraint leads to an unsystematic account of the physical. The NFM constraint is motivated by […] intuitions to the effect that physicalism would be falsified if there turned out to be fundamentally mental entities. But intuitively, physicalism would also be falsified if we were to find that entities at relatively low orders of constitutional complexity were moral

> or freely acting agents, or that aesthetic responses involved a new fundamental interaction or force. Similarly (recalling Driesch and Broad) for chemical, biological and other non-mental, seemingly higher-order features of reality. […] So shouldn't those endorsing a physics-based account of the physical impose, in addition to the NFM constraint, no fundamental morality, no fundamental free will, no fundamental aesthetics, no fundamental chemistry, no fundamental biology, and no miraculous powers constraints? But then, the concern goes, the resulting account of the physical will be unsystematic and ad hoc; for what are mentality, morality, aesthetics, chemistry, biology, and miracles supposed to have in common, that rules them out as being physical? (Wilson 2006: 75)

In my 2006, I aimed to avoid such further constraints in a 'divide and conquer' fashion. As regards fundamental chemistry and biology, I said

> Given that chemical and biological features of reality can, in actual fact, be ontologically accounted for in terms of configurations of […] entities that are not themselves chemical or biological (as all parties to the physicalism debates seem generally prepared to agree), there is no need to explicitly rule these out as being […] fundamental […]. (75)

And for the rest, I argued that insofar as each plausibly involves mentality, no constraint beyond the NFM constraint is needed (76).

This divide and conquer strategy still seems to me to work, but in re the potential need for 'no fundamental chemistry' or 'no fundamental biology' constraints, I now think that something more principled can be said—namely, that these constraints are not needed because chemical and biological goings-on, unlike mental goings-on, are essentially such as to be or be features of comparatively compositionally complex phenomena, such that it would make no sense for individual fundamental physical goings-on, which by the definition of physics are comparatively non-complex, to have chemical or biological features. McLaughlin's question made me realize that there is an important difference here as regards the potential threat of non-mental and mental phenomena so far as characterizing the physical is concerned; for while chemical and biological phenomena might be fundamental in being Strongly emergent (since the advent of such emergence is compatible with, and typically involves, compositional complexity), they could not be fundamental in the sense of being or being features of compositionally basic phenomena. Hence it is, perhaps, that no correlates of panpsychism (panchemism, panbiologism) have been advanced for either chemical or biological features of reality.

I turn next to two concerns that McLaughlin raises for my account of Weak emergence. The first has to do with the whether satisfaction of the conditions in my schema for Weak emergence suffices to render Weak emergents physically acceptable (given the physical acceptability of the base level goings-on). McLaughlin thinks not:

> The nomological requirement on Weak emergence is that if a feature *S* Weakly emerges from a physical feature *P*, then *P* is minimally nomologically sufficient for *S*. That condition is compatible with the law linking *S* and *P* being a fundamental law of nature, a law that doesn't hold in virtue of other laws and conditions. […] The existence of fundamental [e.g.,] psychophysical laws is incompatible with physicalism, reductive or non-reductive. […] To avoid this result, the

> condition of cotemporal material dependence must be amended […] to include the requirement that the law linking *S* and *P* not be a fundamental law of nature; it must be a law that holds in virtue of physical laws and physical conditions (284).

I see McLaughlin's point as in a similar vein to a concern raised by Melnyk (2006). In *Metaphysical Emergence* I present the general concern as follows:

> [W]hatever makes it the case that some proper subsets of token powers of a given lower-level physical feature correspond to (instantiated) higher-level features, while other subsets do not do so, had better itself be physically acceptable if the higher-level features are to be physically acceptable; yet satisfaction of the conditions in Weak Emergence is silent on why a given higher-level feature *S* has the distinctive power profile it has, and so is compatible (one might think) with the instantiation of a higher-level feature's being, somehow or other, the outcome of a physically unacceptable process. (106)

One can develop the concern by noting (as I do in Wilson 2010) that the satisfaction of the proper subset of powers condition is frequently associated with the holding of certain lower-level constraints; as Melnyk correctly observes, if the holding of these constraints ensues as a matter of some physically unacceptable process (say, if the constraints hold as a matter of God's will), then the physical acceptability of the higher-level feature would be thereby undercut. In my 2010, I explicitly require that the constraints at issue be a matter just of physical or physically acceptable processes, and in *Metaphysical Emergence* I register that if an amendment to the schema for Weak emergence is needed, it would likely involve explicitly incorporating this sort of requirement (107).

McLaughlin's comment can be seen as developing the concern in a way that does not specifically advert to constraints, by attention to the possibility that emergent and base features are connected by fundamental laws, as makes sense for Strong but not Weak emergence. And here too I would say that there may well be a case for making the sort of amendment McLaughlin suggests, and requiring that any laws holding between base-level and Weak emergent features hold solely in virtue of physical laws and conditions. That said, rather than bifurcating accounts of the cotemporal material dependence condition which at present is common to the schemas of Weak and Strong emergence, I would prefer to insert any such amendment into the autonomy condition on Weak emergence, to the effect of requiring that any constraints *or laws* operative in making it the case that a given feature is associated with only a proper subset of the token powers of the lower-level base feature be constituted or otherwise determined by lower-level physical processes and/or laws.

McLaughlin also raises the concern that satisfaction of the conditions on Weak emergence is not necessary for metaphysical emergence of a physically acceptable variety. In particular, he suggests that on a 'role-functionalist' view taking higher-level states to be second-order functional states "of being in some state or other that has certain causal effects [where] the first-order states that have those effects realize the functional state" need not be understood as imposing the autonomy (proper subset of powers) condition:

> It is open to a role functionalist to maintain that a functional state, a state of being in some state or other that has certain effects, does not itself cause those effects. Its

> realizers do. That's compatible with functional states figuring in causal explanations of the effects in question. But it is incompatible with Weak emergence. (285)

McLaughlin's suggestion here seems to reflect his position that, while role-functionalism "cannot avoid epiphenomenalism" (McLaughlin 2006: 39), this much does not prevent role-functionalists from adopting "a weaker notion of causal relevance" (one not requiring of a causally relevant feature that it actually cause anything) on which it suffices for a feature to be causally relevant that it be causally 'explanatory'—say, by "providing information about the causal history of an action". Here I'll just say that such a weak understanding of causal relevance is too weak to capture the sense in which we want higher-level features to be efficacious—e.g., as entering into seemingly causal special science laws, or as mental causes of our agential behaviours. Relatedly, such a weak notion of relevance seems ripe for reductive or eliminativist treatment of role-functional features in (mere) conceptual or pragmatic terms (per, e.g., Heil 2003). So on the assumption that role-functionalist features are epiphenomenal, that they don't satisfy the conditions for Weak emergence doesn't pose a problem for my view. That said, it seems to me that role-functionalists can resist the charge of epiphenomenalism, and more specifically can maintain that such properties satisfy the conditions in Weak emergence, for reasons I set out in my book (Wilson 2021: 59–60).

Finally, I turn to McLaughlin's concern that Strong emergence, understood (as on my preferred implementation) as involving a novel fundamental interaction, is incompatible with current physics—in particular, with quantum field theory (QFT), which aims to unify quantum mechanics and special relativity, and is the foundation of the standard model of fundamental particle physics:

> In the field dynamics of quantum field theory, interactions are *local*. They are local in that fields directly interact with other fields only at spacetime points. That is to say, the dynamics of each field at any spacetime point are directly influenced only by the values and derivatives of the other fields at that same point, and not by anything happening elsewhere. That fundamental interactions are local is inextricably baked into the theory. Quantum field theory could, for instance, accommodate new kinds of particles and new kinds of fundamental forces. But the discovery of fundamental configurational interactions would refute the theory. It thus isn't just that quantum field theory doesn't now posit fundamental configurational interactions, it cannot countenance them. Such direct fundamental interactions would involve whole regions of spacetime. That is incompatible with relativity theory. (288)

More specifically, McLaughlin goes on, the enormous success of QFT defeats the considerations I offer for thinking that there is libertarian free will (to wit: that we have direct experience of ourselves as choosing, and that there are presently no good reasons for thinking that we cannot take this experience at realistic face value):

> Quantum field theory has been enormously successful in its regime of applicability, and […] human brains fall well within that regime. The truly enormous empirical support quantum field theory enjoys soundly defeats any intuitions we might have about there being a fundamental force of will. (288)

I offer four lines of response to McLaughlin's objection.

First, it is incorrect that the supposition that fundamental interactions are local, in the sense that fields directly interact only at points, is "inextricably baked into" QFT.[22] To be sure, standard quantum field theory textbooks often claim that interactions are local in this sense, but (as claims in textbook presentations of physical theories often are) this claim is a gloss, which upon closer examination is metaphysically, theoretically, and historically inaccurate.

The usual gloss is metaphysically inaccurate—or at least, metaphysically suspect. To start, field operators are not definable at points unless the theory is fully regulated (rendered non-divergent) in the UV regime. In continuum QFT, field operators must be treated as operator-valued distributions—i.e., one only gets an operator by integrating the distribution against a test function with support on a compact region (i.e., by averaging the field values in a small region around the point), which results in a field observable that is not even gauge invariant. The metaphysical picture encoded in this procedure is murky, and if anything seems to suggest that fields interact not at points, but rather in the compact vicinity of points.[23] Relatedly, the usual means of dealing with UV divergence in local QFT results in a QFT which is an 'effective' field theory, the import of which is precisely to gloss over what exactly is happening at the small-scale limit. Physicists have identified tools (most saliently: renormalization strategies) enabling QFT to be useful for capturing the long distance physics while allowing us to remain agnostic about the short distance physics. But given this understanding of effective QFT, it's clear that there are lots of ways the short distance physics could be. Indeed, there is nothing in QFT itself qua effective theory that demands that what lies below the limit of applicability is even a quantum field theory, much less one that is local (or nonlocal)![24]

The usual gloss is also theoretically and historically inaccurate, since as it happens attention to nonlocal QFT goes back at least to the 1940's and is alive and well today. As Tomboulis (2015) recently put it:

> Nonlocal field theories is a subject with long, albeit spotty, history. Despite the success of perturbative renormalization in QED in the late forties, the idea that local interactions may be a low energy approximation to fundamental underlying nonlocality of interactions continued to be prominent in the fifties and the subject of many investigations [1].[25] Subsequently, nonlocality was considered mostly in

[22] Thanks to Michael Miller and Patrick Fraser for helpful discussion here.

[23] See also the discussion of the 'localization problem' in Saunders 1992.

[24] This is an epistemic point. Interestingly, however, certain metaphysical readings of the effectiveness at issue (say, as involving a lower limit to the precision of the field values, per Miller forthcoming) might also undercut the claim that interactions in QFT occur at points in a continuum.

[25] "[1] R.P. Feynman, Phys. Rev. 74, 939 (1948); A. Pais and G. E. Uhlenbeck, Phys. Rev. 79, 145 (1950); P. Kristensen and C. Møller, Dan. Mat. Fys. Medd. 27, no. 7 (1952); W. Pauli, Nuovo Cimente, 10, 648 (1953); M. Ebel, Dan. Mat. Fys. Medd. 29, no. 2 (1954); M. Chretien and R. E. Peierls, Nuovo Cimento 10, 668 (1953); M Cretien and R. Peierls, Proc. R. Soc. London A223, 468 (1954); C. Hayashi, Prog. Theor. Phys. 10, 533 (1953); ibid, 11, 226 (1954); N. Shono and N. Oda, Prog. Theor. Phys. 8, 28 (1952); F. Bopp, Ann. d. Physik, 42, 573 (1942); H. Mc Manus, Proc. R. Soc. London A195, 323 (1948); G. Wataghin, Z. Phys. 86, 92 (1934)" (26).

> the context of axiomatic field theory [2].[26] In more recent years it has attracted renewed interest in connection with nonlocal theories of gravity [3] - [9],[27] as well as the nonlocality of string field theory vertices and various nonlocal models in cosmology and other areas, see [10][28] and extensive reference list therein. (2)

Others advancing versions of nonlocal QFT include Nobel laureate H. Yukawa,[29] K. Namsrai,[30] G. Fleming,[31] M. Moffat,[32] and R. Landry and J. Moffat.[33] It's clear, then, that physicists do not see the locality of interactions as "inextricably baked into QFT".

There's good reason why nonlocal QFT is of perennial interest as an alternative research program to local QFT. It's not just that local QFT is subject to UV divergence, though that is part of what drives physicists to look elsewhere. As Fleming (1987) observes, the original and continuing motivation for exploring nonlocal QFT reflects concerns "over the internal consistency of a theory requiring infinite renormalization and the long-standing recognition that local interactions generate that requirement". As above, getting any predictions out of QFT requires adopting perturbative methods involving expansions which, unless arbitrarily cut off, give rise to infinities. To be sure, "at the level of comparing renormalized perturbation theory calculations with experiment …[t]he methods work wonderfully!" Still …

> [T]hrough all these years since Dyson, Feynmann, and Schwinger formulated renormalization theory, it has never shed its fundamentally *ad hoc* character. It remains a recipe for extracting finite results from an infinity-plagued formalism by cancelling the infinities against one another systematically. What is wanted is a formulation of non-trivial interacting QFT that never encounters the infinities in the first place. (Fleming 1987: 98–9)

[26] "M. Meyman, Sov. Phys. JETP 20, 1320 (1965); V. Efimov, Com. Math. Phys. 5, 42 (1967); ibid, 7, 138 (1968); M. Z. Iofa and V. Ya. Fainberg, Theor. Mat. Fiz. 1, 187 (1969); M. Z. Iofa and V. Ya. Fainberg,, Sov. Phys. JETP 29, 880 (1969); V. Ya. Fainberg and M. A. Soloniev, Ann. Phys. 113, 421 (1978); V. Ya. Feinberg and M. A. Soloviev, Theor. Math. Phys. 93, 1438 (1992)" (26–27).

[27] "E. T. Tomboulis, arXiv:hep-th/9702146; [4] T. Biswas, E. Gerwick, T. Koivisto and A. Mazumdar, Phys. Rev. Lett. 108, 031101 (2012) [arXiv:1110.5249]; [5] T. Biswas, A. Conroy, A. S. Koshelev and A. Mazumdar, Class. Quant. Grav. 31, 015022 (2014) [arXiv:1308.2319]; [6] L. Modesto, Phys. Rev. D 86, 044005 (2012); [7] L. Modesto, Astron. Rev. 8.2, 4 (2013) [arXiv:11202.3151]; L. Modesto, arXiv:1402.6795[hep-th]]; F. Briscese, L. Modesto and S. Tsujikawa, Phys. Rev. D 89, 024029 (2014) [arXiv:1308.1413]; G. Calcagni and L. Modesto, Phys. Rev. D 91, 124059 (12015) [arXiv:1404.2137 [hep-th]; L. Modesto and L. Rachwal, Nucl. Phys. B889, 228 (2014) [arXiv:1407.8036]. [8] M. Isi, J. Mureika and P. Nocolini, JHEP 1311:139 (2013) [arXiv:1310.8153 [hep-th]]. [9] V. P. Frolov, arXiv:1505.00492; V. P. Frolov, A. Zelnikov and T. de Paula Netto, arXiv:1504.00412" (27).

[28] "N. Barnaby and N. Kamran, JHEP 0802, 008 (2008)" (27).

[29] See in particular Yukawa 1950a and 1950b.

[30] See, e.g., Namsrai 1986.

[31] See, e.g., Fleming 1987.

[32] See, e.g., Moffat 1990.

[33] See Landry and Moffat (forthcoming).

The deeper motivation for exploring nonlocal QFT is that the assumption of locality itself underlies UV divergence. As Tomboulis (2015) puts it:

> It has long been realized, more or less explicitly, that UV finiteness (or at least superrenormalizability in the presence of gauge interactions) can be achieved by nonlocal interactions. (2)

Of course, UV finiteness isn't the only theoretical desideratum. In addition, theorists want QFT to satisfy unitarity and causality, in a way compatible with relativity. Tomboulis goes on:

> [On nonlocal QFT], unitarity can be preserved, at least perturbatively, provided appropriate analyticity conditions can be imposed on the nonlocal interactions. Causality, however, is a central concern whose investigation has remained woefully inadequate, both in the classical theory, where it is inexorably connected with the mathematically proper formulation of the initial value problem (IVP), and in the quantum theory. (2)

In any case, many nonlocal versions of QFT claim to avoid UV divergence while accommodating both unitarity and causality. For example, Namsrai (1986) constructs "a nonlocal theory of quantized fields by means of the hypothesis of *spacetime stochasticity*", and Fleming (1987) formulates a nonlocal QFT involving spacelike hyperplanes:[34]

> Hyperplane dependence of the dynamical variables of quantum theory, and consequently, their eigenvectors, is the minimal generalization of the concept of time dependence that is required to establish a manifestly Lorenz covariant formalism. […] The reason that hyperplane dependence has not previously become a prominent conceptual tool of theoretical physics [reflects that] contemporary fundamental theories of many-particle systems are expressed in terms of basic quantized fields that are themselves associated with simple points of space-time. [But this line of thought] may be unnecessarily restrictive. The experience my students and I have gained, in exploring the possibilities, allowed for interactions of particles with external potentials when hyperplane dependence is explicitly incorporated into the formalism, and suggests the possibility that consistent Lorentz-invariant quantum field theories with nonlocal interactions may be possible if the fields are hyperplane-dependent. I will suggest below a model of such a theory. (97–8).

In discussing Fleming's view, Saunders (1992: 379) suggests that a relaxing of the demand for local covariance, to be replaced in particular by the weaker requirement of hyperplane dependent covariance, may well be "all but inevitable". Yet more recently, Landry and Moffat (forthcoming) say:

> We discuss the nonlocal nature of quantum mechanics and the link with relativistic quantum mechanics such as formulated by quantum field theory. We use here a nonlocal quantum field theory (NLQFT) which is finite, satisfies Poincaré

[34] A spacelike hyperplane is a three-dimensional, metrically flat section of the flat Minkowski space-time continuum, such that any two points in the hyperplane are separated by a spacelike interval, and such that for any such hyperplane, there is an inertial frame of reference in which all the points of the hyperplane are simultaneous, and all points simultaneous with any point of the hyperplane are in the hyperplane.

> invariance, unitarity and microscopic causality. This nonlocal quantum field theory associates infinite derivative entire functions with propagators and vertices. We focus on proving causality and discussing its importance when constructing a relativistic field theory. […] The result is free of UV divergences and we recover the area law.

Suffice to say that nonlocal QFT is a research program with a long history that people are still actively pursuing.[35]

Third, it's not clear that any Strong emergence there might be would violate microcausality. To start, note that any demand for locality in QFT had better be compatible with entanglement; and indeed it is, since the locality characteristic of QFT is one supposed to preserve "microcausality", whereby no causal influences can travel faster than the speed of light. Entanglement phenomena don't violate microcausality, and so don't violate locality in that sense; rather, they violate separability, according to which the wave-function for the system as a whole is factorizable as a product of wave-functions for the system's parts. In this sense, entangled systems are irreducibly holistic, with a common spin (no pun intended) being that entangled particles are not really distinct; hence it is that for one entangled particle to "influence" another does not require faster-than-light (or any) causal connections. (Or so the story goes.) Now return to Strongly emergent phenomena. These are often characterized in terms evocative of failures of separability: a Strongly emergent feature is one which cannot be factored or otherwise reduced to features of its parts. Moreover, the failure of reduction here is one according to which a Strongly emergent feature is holistic, in arising (in this context) under conditions of compositional complexity, with a common spin on such features being that they render the system that has them a unified whole, whose parts are not really distinct. These similarities suggest that on the face of it, Strongly emergent features, like entangled systems, would violate separability, not microcausality.

That said, in my book I argue that entanglement phenomena are not in general clear cases of Strong emergence, since the failure of reduction might be understood as involving Weak emergence from a spatiotemporally extended dependence base. Strong emergence, on my view, involves a fundamentally novel power, which in turn (on my preferred implementation, and as motivated by the case of the weak nuclear interaction; see my 2002 and 2021) involves a novel fundamental interaction which comes into play only at certain levels of compositional complexity. How would this work? Well, whatever is going on here, it won't be a matter of instantaneous causal influences. Rather, on the usual assumption that fundamental interactions are associated with fields, Strong emergence would involve a new fundamental field (or fields) coming into play, which would presumably interact with other fields/interactions in operation, just as standardly posited fields/interactions do. How, exactly, and what theoretical and empirical consequences this would have, would sensitively depend on the nature of the interaction between the standard fields and the new field(s), which as in the case of standard fields/interactions would be an a posteriori, empirical matter. For present purposes it suffices to note that there is no in-principle barrier to understanding Strong emergence in this way,

[35] It may also be worth noting that, as Weinberg (1997) observes, QFT as standardly formulated is not fully either nonlocal or Lorentz invariant: "there are complications when you have things like mass zero, spin one particles for example; in this case you don't really have a fully Lorentz invariant Hamiltonian density, or even one that is completely local" (7).

and indeed (again, see my discussions of the weak nuclear interaction) there is some historical precedent for doing so.

Fourth, though for the reasons above there's no clear conflict between Strong emergence and QFT, it's worth noting that McLaughlin's claim (following Carroll 2021) that QFT "has been enormously successful in its regime of applicability, and […] human brains fall well within that regime" (288) involves a massive and to my mind unjustified extrapolation. As Carroll himself observes,

> Particle-physics experiments typically examine the interactions of just a few particles at a time, so new physical laws that only kick in for complex agglomerations of particles are not necessarily ruled out by data we currently have (2021: 28).

In that case, though, why think that "particles obey the same equations whether they are inside a rock or inside a human brain" (27), contra applications of Strong emergence to mental phenomena such as (in my book) libertarian free will? Here Carroll appeals to the status of QFT as an effective theory targeting low-energy states, which can be interpreted as collections of interacting particles. Insofar as human beings, like rocks, can (under decomposition) be thought of as such collections, they fall in the regime of applicability of QFT. But the true measure of a theory's "applicability" is predictability, not the fact that, as Carroll puts it, the theory "is meant to be accurate" (18) for phenomena in some or other energy regime. And QFT provides no predictive basis for any human behaviour, unlike the remarkably successful predictions we make through understanding our own and others' mental states. On the face of it, then, McLaughlin's extrapolation, like Carroll's, requires assuming that there are no new fundamental configurational interactions or laws—that, as a synchronic variation on Hume's problem of induction, the physical laws of nature "will continue the same".[36] But like Hume's problem, that assumption builds in what the argument from QFT is supposed to show.

For the various reasons above, I conclude that attention to QFT poses no in-principle difficulty for Strong emergence. But no doubt there is more to say here, and I thank McLaughlin (and Carroll) for raising this important question to salience.

8. Replies to Paolini Paoletti

In his contribution, Paolini Paoletti raises two questions pertaining to the metaphysics of properties, as potentially relevant to my schema for Weak metaphysical emergence. The first question presupposes (correctly, in my view) that in general,

[36] Carroll also says that "if there are additional particles and forces, they interact too weakly with the known fields to exert any influence on human behavior; otherwise they would have already been detected in experiments" (2021: 18). But again, as Carroll notes, the experiments that have been so far conducted are limited to examining "the interactions of just a few particles at a time" (28), far below the complexity at which, e.g., Strongly emergent mental features are supposed to exist or be instantiated. To be sure, if Strong emergence involves the coming into play of a new fundamental interaction, then once such an interaction is on the scene it could (in principle) have theoretical or empirical consequences for interactions involving systems at lower levels of complexity; but whether this would be the case would be an empirical matter.

not every proper subset of powers associated with a given physical feature *P* is associated with a Weakly emergent feature. In that case, one can ask:

(1) What makes it the case that a given proper subset of powers associated with a given lower-level physical feature is associated with a Weakly emergent feature?[37]

The second question presupposes that features can be individuated in a way independently of their powers. In that case, one can ask:

(2) What makes it the case that a given feature *S* is associated, with at least nomological necessity, with a given causal profile?

Paolini Paoletti considers certain candidate answers to these questions, and finds them wanting. He then advances essence-based answers to these questions—but, he maintains, an essence-based approach is in tension with the supposition that "everything whatsoever is physical or fully depends on the physical" (311), such that Weak emergence turns out to be "not so weak", after all.

Now, as Paolini Paoletti notes, I don't aim in my book to answer either question. In re the first question: in my book and elsewhere I take for granted what I call the prima facie appearances of metaphysical emergence in the sciences and in ordinary experience, as coupling dependence with ontological and causal autonomy; and then I argue that in various cases we can make sense of these prima facie appearances—most commonly, as satisfying the conditions in the schema for Weak emergence. In cases of broadly scientific properties, for example: what explains why scientists have posited certain higher-level scientific properties as having certain subsets of powers, as is reflected in these properties' entering into certain special-science laws? I discuss certain broadly empirical motivations which seem to be operative in some cases (upon which I'll expand below), but ultimately I take this to be a question for the (natural and social) scientists. My job, as I see it, is just to show that one can make metaphysical good sense of such posits. And in re the second question: as I further discuss below, this question arises only for those holding certain metaphysical views of features (properties and the like) and powers—in particular, those who think that features can be individuated independently of their powers—in the usual case, via a quiddity or primitive identity, which can then be somehow associated or not associated with certain powers. My own view is that there is no reason to think that features of the sort under discussion in my book are associated with quiddities or any other kind of non-causal aspects, in which case the second question doesn't arise, though I also argue in my book that the viability of the schemas for emergence is neutral on whether features are associated with quiddities.

All this said, one way to read the intended import of Paolini Paoletti's remarks is that if one *does* attend to these questions, one will see that they interestingly bear on how Weak emergence should best be understood, and on whether Weak emergence (properly understood) can provide a satisfactory basis for nonreductive physicalism. So in what follows I start by arguing that answers to the first question are plausibly both diverse and empirical, as are answers to the second question as it arises for those accepting quiddities or other non-causal aspects of properties. I'll then follow up by offering reasons to reject a thesis that enters into Paolini Paoletti's critical assessment of certain strategies for answering his

[37] I phrase this and the second question in terms of "what makes it the case" that *P* as opposed to why *P*, in order to sidestep cases where $\neg P$.

questions—namely, Sider's principle of 'Purity' (see Sider 2011: 126–132), according to which the constituents of fundamental facts must themselves be fundamental. Finally, I raise some concerns with Paolini Paoletti's positive "essence-based" answers to the questions, and relatedly, with his claim that if (as on his preferred answers), a higher-level feature and its causal role are in some sense mutually essentially dependent, this poses a problem for physicalism understood as requiring that "everything […] fully depends on the physical" (311).

To start, then: what makes it the case that a given proper subset of lower-level physical powers is associated with a higher-level Weakly emergent feature? This is a question of general interest, whatever one's metaphysics of properties, at least for those who accept that there is or may be Weak emergence. In my book I discuss some of the considerations motivating scientific posits of certain higher-level features having certain causal profiles. One common answer, which I discuss in Ch. 3 in presenting my DOF-based approach to Weak emergence, adverts to there being certain conditions or associated constraints present at the lower level, which serve to eliminate certain microphysical degrees of freedom as required for characterizing the law-governed properties and behaviour of the higher-level feature (which elimination in DOF in turn operates to eliminate certain powers as had by the feature). A different but related consideration, which I discuss in Ch. 5 in motivating the claim that certain complex systems are Weakly emergent, adverts to the suitability for a given complex system to be modeled by the Renormalization Group Method, which in turn reflects that the system ceases to have a preferred length scale—which again serves to eliminate certain lower-level physical DOF and associated powers. So here we have one sort of broadly (lower-level constraint-based) empirical answer to the first question, which Paolini Paoletti considers under the heading of my 'physicalistic solution,' and which he takes to be successful—in particular, "fully compatible with all versions of physicalism" (308), as far as it goes.

As Paolini Paoletti observes, however, my DOF-based account is only presented as a sufficient implementation of the schema for Weak emergence, and so won't work by way of a general answer to his questions; and indeed, as I clarify in my reply to Emery (this volume), other cases of Weak emergence are not clearly ones involving an elimination in DOF; so in these other cases a different answer to the first question might be operative. For a determinable-based implementation of Weak emergence of the sort that seems promising as applied to perceptual mental states, answering the first question would involve exploring why a given determinate has the determinables it does, which would require (among other things) attention to the determination dimensions of the determinate (see Funkhouser 2006). For a functional realization-based implementation of Weak emergence of the sort that seems promising as applied to artifactual features, answering the first question would involve exploring why certain functional roles are salient in our social economy. So here we have different sorts of answers to the first question, but so far as I can tell, these will also be broadly empirical, in depending on complex, broadly contingent facts. As such, even granting the general interest of the first question, I don't see any reason to think that it will have a single or unified answer, much less a single or unified metaphysical answer, of the sort that Paolini Paoletti appears to be seeking.

What about the second question, of what makes it the case that a given feature *S* is associated, with at least nomological necessity, with a given causal profile? Again, it seems to me that this question arises only for those who think that

features can be individuated independently of their powers via quiddities or primitive identities. Paolini Paoletti seems to take such a view for granted in his attempt to answer this question; hence, e.g., in considering whether the connection of a given causal profile to a given property is primitive, he says, "To make sense of this situation from an ontological standpoint, we may hold that there is some irreducible relation *R* that links *S* (and only *S*) with its causal role (and only with it)" (306). He rejects this primitivist answer, for reasons I'll discuss down the line, but the terms of the solution, like the question itself, presuppose that one may refer to a feature in some way independent of its powers—which those rejecting quiddities or the like will deny. Paolini Paoletti suggests that even someone not endorsing quiddities will have to answer a version of the second question. Hence he says of a non-quiddistic view on which properties are mere bundles of token powers that, "one would still need to explain why only certain bundles of token powers (and not others) seem to 'give rise to' or 'be legitimately describable as' token features" (308). But first, one may reject quiddities without embracing a bundle theory (which on the face of it reifies powers in a way that I would resist); one may rather simply think of properties in what I think of as metaphysically adverbial terms, as ways things are.

In any case, the (second) question as directed at the non-quidditist of whatever variety isn't the same as that directed at the quidditist. The question for the non-quidditist can be understood in two ways, depending on whether it is asked against a backdrop assumption of there being lower-level physical features associated with specific causal profiles. If so, then the question collapses into the first question—i.e., what makes it the case that a given subset of physical powers corresponds to a genuine feature? If not—if the question is more generally asking which collections of powers or "ways things are" correspond to genuine properties—the question collapses into the question "Which properties exist?" That's an interesting question, to which whole fields are devoted—but not one that any individual metaphysician has the burden of answering.

Putting my own inclinations aside, it seems to me that proponents of quidditistic accounts of properties typically suppose that the answer to Paolini Paoletti's second question is an empirical matter, even if they disagree over details. Hence, for example, Lewis (1986) supposes that what powers are associated with which (intrinsic, categorical) properties is a matter of the distribution of those properties in the Humean mosaic, which metaphysically contingent distribution determines the laws of nature at the world; and Armstrong (1983) supposes that what powers are associated with which universals is a matter of which metaphysically contingent relations of nomological necessitation hold at the world. Either way, answers to Paolini Paoletti's second question will be both diverse (depending on further commitments of the individual quidditist) and ultimately advert to certain contingent empirical facts.

I want to turn now to a thesis that shows up in Paolini Paoletti's assessment of a primitivist response to the first and second questions. Focusing on a specific instance of the first question, he says "suppose that we claim that it is a primitive and inexplicable fact of the matter that the proper subset made of $p1$, $p2$ and $p3$ (i.e., the causal role of *S*) is the only one that is associated with a higher-level token feature" (305). He goes on:

> [T]hat *R* holds between *S* and its causal role is an irreducible fact of the matter. Thus, it is a fundamental fact. Moreover, this fact constitutively includes a non-

> physical token feature such as *S*. Thus, there are fundamental facts with non-physical token features such as *S*. The constituents of fundamental facts are fundamental [following Sider 2011]. Therefore, non-physical token features such as *S* are fundamental. This conclusion may be hard to swallow for physicalists. (306)

Clearly it would be problematic for physicalists were a given higher-level feature, that was supposed to be Weakly emergent and so (though physically acceptable) not identical to any physical feature, turned out to be fundamental; for physicalists of any variety maintain that lower-level physical goings-on are the only fundamenta there are. Now, as above, I don't think there's any pressure here to embrace primitivism about the first or second questions, since each admits of diverse, broadly empirical answers. That said, Paolini Paoletti's remarks offer me an opportunity[38] to rail against Sider's purity principle (for short: 'Purity')—again, according to which the constituents of fundamental facts must themselves be fundamental.

In brief: I see no reason to accept Purity, and on the contrary good reason not to do so. For the fundamental goings-on—whether these be facts, states of affairs, or some other constituents of reality—are (if nothing else) required to serve as a suitable basis for all of reality, including any non-fundamenta there might be. Everyone agrees on this much, whatever the further details of their preferred account of what makes it the case that some goings-on at a world are fundamental at that world.[39] Hence it is that characterizations of fundamentality often start with the familiar "All God had to do" heuristic, according to which the fundamental goings-on are all God had to create in order to create the world as a whole. But if the world as a whole flows, one way or another, from just the fundamenta, then far from supposing that the fundamenta cannot contain or encode reference to non-fundamenta, it seems on the contrary that the fundamenta must contain or encode reference to non-fundamenta, for otherwise it is opaque how they could bring the non-fundamenta in their wake. Hence Purity is false. A better characterization of fundamental facts, it seems to me, is one according to which a fundamental fact must contain at least one fundamental entity or feature as a constituent; but that's compatible with fundamental facts' containing non-fundamenta as well. In any case, given that Purity is (to my mind: clearly) false, Paolini Paoletti's rejection of primitivist answers to his questions will have to rely on considerations (e.g., parsimony concerns of the sort he discusses) other than their leading to a supposed violation of Purity.

I now want to move on to Paolini Paoletti's preferred essentialist approach to his two questions. He maintains:

[38] Or another opportunity: see Wilson 2018 for an initial salvo.

[39] Among the usual suspects here are independence-based accounts (what makes it the case that some goings-on are fundamental is that they are independent of all else; see Schaffer 2009, Bennett 2017), complete minimal-basis accounts (what makes it the case that some goings-on are fundamental is that they are part of a minimal collection of goings-on which serve as a basis for all else; see Tahko 2018), and primitivist accounts (what makes it the case that some goings-on are fundamental is a primitive matter, not metaphysically analyzable in any other terms—a view which is, by the way, compatible with it being necessary or even essential to the fundamenta at a world that they enter into a basis for all else at the world; see Fine 2001, Wilson 2014 and forthcoming).

> [T]he best way to answer questions (1) and (2) consists in embracing something akin to 'grounding categoricalism,' i.e., the doctrine according to which the causal roles of categorical properties are somehow grounded on those very properties (see, among others, Tugby 2012, 2020, 2022, Yates 2018, Kimpton-Nye 2021, Paoletti 2021). In Paoletti (2021), I have defended the following form of grounding categoricalism: by virtue of its own essence, the causal role C of a categorical property P (i) is the causative role of P, so that it essentially depends (also) on P, (ii) it depends for its origins on P (i.e., it starts to exist as a causal role thanks to P or thanks to the instantiation of P) and (iii) it depends for its continuing to exist (also) on P (i.e., it continues to exist also or only thanks to P or to the instantiation of P). This entails that, as a matter of necessity, the existence of C implies the existence of P: necessarily, C cannot exist without P. And it also entails that, as a matter of necessity, C is the causal role of P and of no other property distinct from P. (308–309)

Here by the "essence" of an entity, Paolini Paoletti means "what that entity nonderivatively is (or could be) in all possible circumstances" (309).

In what follows I'll register certain concerns about an essence-based approach to the questions at issue, and with Paolini Paoletti's claim that such an approach has substantive implications for our understanding of physicalism, and more specifically, of Weak emergence.

First, Paolini Paoletti claims that grounding categoricalism provides attractive answers to the questions he has posed, but I don't see that this is so. Taking properties to be essentially such as to have or be otherwise associated with certain causal roles, which as it happens are comprised of a specific proper subset of lower-level physical powers (per the schema for Weak emergence), certainly provides a mechanism whereby a property and its causal profile go hand in hand, but it doesn't illuminate why (as regards the first question) only certain subsets of lower-level powers are associated with higher-level Weakly emergent properties, or why (as regards the second) a given property is essentially such to have *these* powers, either as a matter of nomological or metaphysical necessity. Again, I'm inclined to think that these questions admit of empirical answers, but Paolini Paoletti seemed to want more—namely, some metaphysical account of why certain collections of lower-level powers, but not others, are associated with some or other feature (the first question), and moreover with a specific feature (the second question). I don't see how grounding categorialism provides such an account, since that causal profiles are essentially tied to properties doesn't tell you what causal profiles and associated properties there are. Rather, grounding categoricalism introduces a slew of new questions, including: which essences are there? and why is a given essence associated with this causal role in this world (worse: at this time) and that causal role in that world (that time)?

If the answers to these questions turn out to be "it's primitive", then Paolini Paoletti's (remaining) concerns with primitivist answers to his question attach also to his account. Now, Paolini Paoletti claims that with his essence-based solution, "we avoid introducing primitive and sui generis connections" (310) between features and causal profiles, but to my mind an appeal to causal profiles as "grounded" in essences just pushes, and indeed multiplies, the primitivist bump(s) in the rug. Paolini Paoletti asserts that the grounding connections are not primitive, since "internal", but even granting that internal relations do not introduce primitive posits, the claim that the relation between essences and causal profiles

is internal doesn't establish this much; since no handle on the relation at issue has been provided sufficient unto showing that the relation is in fact internal.[40]

These considerations provide, in my view, good reason to stick with the usual array of empirical considerations offered by scientists and philosophers as motivating there being these special science features and associated powers/laws, and not others, which methodological strategy provides a generally explanatory and comparatively parsimonious basis for answering Paolini Paoletti's questions (to the extent that one feels pressure to do so, as a consequence of one's independent commitments---e.g., to a quidditistic conception of properties). Here it is also worth noting that one can deny Paolini Paoletti's claim that quiddities are motivated as answering his questions, since as previously discussed, there are available broadly empirical answers to the first question, and the second question doesn't arise unless one posits something like quiddities—in which case a purported need to answer his second question doesn't provide independent reason to posit quiddities.

Second and finally, even if it turns out that properties and their powers are essentially mutually dependent, I don't see that there is a deep problem for physicalism here. Physicalism is the view that all broadly scientific goings-on are "nothing over and above" lower-level physical goings-on, in the way that reductive versions of physicalism (appealing to identity) or non-reductive versions of physicalism (appealing to functional or other forms of realization, the key features of which are encoded in the schema for Weak emergence) aim to capture. It isn't any part of the physicalist project to maintain that mathematical or metaphysical features—e.g., the property of being prime, the relation between a universal and its instantiation, or (if such there be) the relation between a feature and its causal profile—are in any way nothing over and above or completely dependent on lower-level physical goings-on. So even if one is inclined to follow Paolini Paoletti in taking an essence-based approach to the questions he has raised, this in itself poses no tension with physicalism, or so it seems to me.

References

Albert, D., 2013. Wave function realism. *In*: A. Ney and D. Albert, eds. *The wave function: Essays on the metaphysics of quantum mechanics*. New York: Oxford University Press, 52–57.

Armstrong, D.M., 1978. *Universals and scientific realism, Vol i: Nominalism and Realism*. Cambridge: Cambridge University Press.

Armstrong, D.M., 1983. *What is a law of nature?* Cambridge: Cambridge University Press.

Baker, D.J., 2020. Knox's inertial spacetime functionalism. *Synthese*, 199 (S2), 277–298. https://doi.org/10.1007/s11229-020-02598-z.

[40] In addition, Paolini Paoletti appeals to diverse, but unspecified, dependence relations in order to accommodate a purported "circle of dependence" between features and profiles, in which case his approach appears to involve (in addition to its being primitive what essences of features there are, and primitive that certain causal profiles depend on such essences) a third primitive component, tracking that the features at issue depend on the associated profiles.

Baron, S. 2019. The curious case of spacetime emergence. *Philosophical studies*, 177, 2207–2226. https://doi.org/10.1007/s11098-019-01306-z.

Baron, S. and Le Bihan, B., 2022. Composing spacetime. *Journal of philosophy*, 119, 33–54. https://doi.org/10.5840/jphil202211912.

Baron, S. and Miller, K., 2014. Causation in a timeless world. *Synthese*, 191, 2867–2886. https://doi.org/10.1007/s11229-014-0427-0.

Baysan, U., 2016. An argument for power inheritance. *Philosophical Quarterly*, 66, 383–390.

Baysan, U. and Wilson, J.M., 2017. Must strong emergence collapse?. *Philosophica*, 91, 49–104.

Bellazzi, F., Biochemical functions. *Forthcoming. British journal for philosophy of science.*

Bennett, K., 2003. Why the exclusion problem seems intractable and how, just maybe, to tract it. *Noûs*, 37, 471–497.

Bennett, K., 2008. Exclusion again. *In*: J. Hohwy and J. Kallestrup, eds. *Being reduced: New essays on reduction, explanation, and causation*. Oxford: Oxford University Press, 280–307.

Bennett, K., 2017. *Making things up*. Oxford: Oxford University Press.

Bohn, E., 2009. Composition as identity: A study in ontology and philosophical logic. Thesis (PhD). University of Massachusetts.

Calosi, C., 2016a. Composition, identity, and emergence. *Logic and Logical Philosophy*, 25, 429–443. https://doi.org/10.12775/llp.2016.010.

Calosi, C., 2016b. Composition is identity and mereological nihilism. *Philosophical Quarterly*, 66, 219–235. https://doi.org/10.1093/pq/pqv109.

Calosi, C. and Giordani, A., Atoms, combs, syllables and organisms [unpublished manuscript].

Calosi, C., Universalism and extensionalism revisited. [unpublished manuscript].

Calosi, C. and Wilson, J.M., 2018. Quantum metaphysical indeterminacy. *Philosophical Studies*, 176, 2599–2627.

Cameron, R., 2007. The contingency of composition. *Philosophical Studies*, 136, 99–121.

Canavotto, I. and Giordani, A., 2020. An extensional mereology for structured entities. *Erkenntnis*, 87, 2343–2373. https://doi.org/10.1007/s10670-020-00305-5.

Carroll, S.M., 2021. Consciousness and the laws of physics. *Journal of Consciousness Studies*, 28, 16–31.

Chalmers, D.J., 2021. Finding space in a nonspatial world. *In*: C. Wüthrich, B. Le Bihan, and N. Huggett, eds. *Philosophy beyond spacetime*. Oxford: Oxford University Press.

Clapp, L., 2001. Disjunctive properties: Multiple realizations. *Journal of Philosophy*, 98, 111–136.

Cornell, D.M., 2017. Mereological nihilism and the problem of emergence. *American Philosophical Quarterly*, 54, 77-87.

Cotnoir, A. and Varzi, A., 2021. *Mereology*. Oxford: Oxford University Press.

Crowther, K., 2018. Inter-theory relations in quantum gravity: Correspondence, reduction and emergence. *Studies in History and Philosophy of Modern Physics*, 63, 74–85. https://doi.org/10.1016/j.shpsb.2017.12.002.

Crowther, K., 2022. Spacetime emergence: Collapsing the distinction between content and context? *In*: S. Wuppuluri and I. Stewart, eds. *From electrons to elephants and elections: Saga of content and context*. Cham, Switzerland: Springer, 379–402.

Fine, K., 1975. Critical notice of Lewis, *Counterfactuals*. *Mind*, 84, 451–58.

Fine, K., 2001. The question of realism. *Philosophers' imprint*, 1, 1–30.

Fine, K., 2010. Towards a theory of part. *Journal of philosophy*, 107, 559–589.

Fleming, G., 1987. Hyperplane dependent quantized fields and Lorentz invariance. *In*: H. Brown and R. Harre, eds. *Philosophical Foundations of Quantum Field Theory*, Oxford: Clarendon Press, 93–115.

Funkhouser, E., 2006. The determinable-determinate relation. *Noûs*, 40, 548–69.

Gillett, C., 2016. *Reduction and emergence in science and philosophy*. Cambridge: Cambridge University Press.

Guay, A. and Sartenaer, O., 2016. A new look at emergence. Or when after is different. *European journal for philosophy of science*, 6, 297–322. https://doi.org/10.1007/s13194-016-0140-6.

Heil, J. 2003. Levels of Reality. *Ratio*, 16, 205–221.

Huggett, N., 2021. Spacetime "emergence". *In*: E. Knox and A. Wilson, eds. *The Routledge Companion to Philosophy of Physics*. New York: Routledge, 374–385.

Humphreys, P., 1997. How properties emerge. *Philosophy of science*, 64, 1–17.

Kim, J., 1989. The myth of nonreductive materialism. *Proceedings and addresses of the American Philosophical Association*, 63, 31–47.

Kim, J., 1998. *Mind in a physical world*. Cambridge: MIT Press.

Kim, J., 2006. Emergence: Core ideas and issues. *Synthese*, 151, 547–59.

Kimpton-Nye, S., 2021. Reconsidering the dispositional essentialist canon. *Philosophical studies*, 178, 3421–3441. https://doi.org/10.1007/s11098-021-01607-2.

Koslicki, K., 2008. *The structure of objects*. Oxford: Oxford University Press.

Lam, V. and Wüthrich, C., 2018. Spacetime is as spacetime does. *Studies in history and philosophy of science Part B: studies inhHistory andphilosophy of modern Physics*, 64, 39–51. https://doi.org/10.1016/j.shpsb.2018.04.003.

Lamb, M., 2015. Characteristics of non-reductive explanations in complex dynamical systems research. Thesis (PhD). University of Cincinnati.

Landry, R. and Moffat, J., forthcoming. Nonlocal quantum field theory and quantum entanglement. *European physical journal*, https://doi.org/10.48550/arXiv.2309.06576.

Lewis, D., 1986. *On the plurality of worlds*. London: Blackwell.

McDaniel, K., 2008. Against composition as identity. *Analysis*, 68 (2), 128–133. https://doi.org/10.1093/analys/68.2.128.

McLaughlin, B., 2006. Is role-functionalism committed to epiphenomenalism? *Journal of consciousness studies*, 13, 39–66.

McLaughlin, B. and Bennett, K., 2018. Supervenience. *In*: E.N. Zalta, ed. *The Stanford encyclopedia of philosophy*, Winter 2018. https://plato.stanford.edu/archives/win2018/entries/supervenience/; Metaphysics Research Lab, Stanford University.

Melnyk, A., 2006. Realization-based formulations of physicalism. *Philosophical Studies*, 131, 127–155.

Miller, M., forthcoming. Worldly imprecision. *Philosophical studies*, [unpublished manuscript].

Moffat, J.W., 1990. Finite nonlocal Gauge field theory. *Phys. Rev. D*, 41, 1177–1184. https://doi.org/10.1103/PhysRevD.41.1177.

Morrison, M., 2012. Emergent physics and micro-ontology'. *Philosophy of science*, 79, 141–166.

Namsrai, K., 1986. *Nonlocal quantum field theory and stochastic quantum mechanics*. Dordrecht: Springer Netherlands.

Ney, A., 2021. *The world in the wave function*. Oxford: Oxford University Press.

Paoletti Paolini, M., 2021. Emergence and structural properties. *Synthese*, 198, 8755–8778. https://doi.org/10.1007/s11229-020-02599-y.

Rovelli, C., 2007. *Quantum gravity*. Cambridge: Cambridge University Press.

Santos, G., Vallejos, G., and Vecchi, D., 2020. A relational-constructionist account of protein macrostructure and function. *Foundations of Chemistry*, 22, 363–382.

Sattig, T., 2015. *The double lives of objects: an essay in the metaphysics of the ordinary world*. Oxford: Oxford University Press.

Saunders, S.W., 1992. Locality, complex numbers, and relativistic quantum theory. *PSA: Proceedings of the biennial meeting of the Philosophy of Science Association*, 365–380.

Schaffer, J., 2009. On what grounds what. *In*: D. Manley, D. Chalmers, and R. Wasserman, eds. *Metametaphysics: New essays on the foundations of ontology*, 347–83. Oxford: Oxford University Press.

Schaffer, J., 2010. Monism: The priority of the whole. *Philosophical review*, 119, 31–76.

Shoemaker, S., 2000/2001. Realization and mental causation. *In*: *Proceedings of the 20th World Congress in Philosophy*, Cambridge: Philosophy Documentation Center, 23–33.

Sider, T., 2011. *Writing the book of the world*. Oxford: Oxford University Press.

Simons, P.M., 1987. *Parts: a study in ontology*. Oxford: Oxford University Press.

Tahko, T.E., 2018. Fundamentality and ontological minimality. *In*: R. Bliss and G. Priest, eds. *Reality and its structure*. Oxford: Oxford University Press, 237–53.

Tallant, J., 2019. Causation in a timeless world? *Inquiry: an interdisciplinary journal of philosophy*, 62, 309–325. https://doi.org/10.1080/0020174x.2018.1446051.

Taylor, E., 2015. Collapsing emergence. *Philosophical Quarterly*, 65, 732–753.

Tomboulis, E.T., 2015. Nonlocal and quasilocal field theories. *Phys. Rev. D.*, https://doi.org/10.1103/PhysRevD.92.125037

Tugby, M., 2012. Rescuing dispositionalism from the ultimate problem: Reply to Barker and Smart. *Analysis*, 72, 723–731. https://doi.org/10.1093/analys/ans112.

Tugby, M., 2020. Grounding theories of powers. *Synthese*, 198 (12), 11187–11216. https://doi.org/10.1007/s11229-020-02781-2.

Tugby, M., 2022. Dispositional realism without dispositional essences. *Synthese*, 200 (3), 1–27. https://doi.org/10.1007/s11229-022-03554-9.

Weinberg, S., 1997. What is quantum field theory, and what did we think it is? https://arxiv.org/abs/hep-th/9702027.

Wilson, J.M., forthcoming. The fundamentality first approach to metaphysical structure. *Australasian Journal of Philosophy*.

Wilson, J.M., under contract. *Fundamentality and metaphysical dependence*. Oxford: Oxford University Press.

Wilson, J.M., 2002. Causal powers, forces, and superdupervenience. *Grazer Philosophische-Studien* 63, 53–78.

Wilson, J.M., 2005. Supervenience-based formulations of physicalism. *Noûs*, 39, 426–459.

Wilson, J.M., 2006. On characterizing the physical. *Philosophical Studies*, 131, 61–99.

Wilson, J.M., 2010. Non-reductive physicalism and degrees of freedom. *British journal for the philosophy of science*, 61, 279–311.

Wilson, J.M., 2013. A determinable-based account of metaphysical indeterminacy. *Inquiry*, 56, 359–385.

Wilson, J.M., 2014. No work for a theory of grounding. *Inquiry*, 57, 1–45.

Wilson, J.M., 2015. Metaphysical emergence: Weak and strong. *In*: T. Bigaj and C. Wüthrich, eds. *Metaphysical emergence in contemporary physics; Poznan Studies in the Philosophy of the Sciences and the Humanities*. Amsterdam/New York: Brill, 251–306.

Wilson, J.M., 2018. Grounding-based formulations of physicalism. *Topoi*, 37, 495–512.

Wilson, J.M., 2021. *Metaphysical emergence*. Oxford: Oxford University Press.

Yates, D., 2018. Inverse functionalism and the individuation of powers. *Synthese*, 195: 4525–4550. https://doi.org/10.1007/s11229-017-1417-9.

Yukawa, H., 1950a. Quantum theory of non-local fields. Part I. Free fields. *Phys. Rev.*, 77 (January): 219–226. https://doi.org/10.1103/PhysRev.77.219.

Yukawa, H., 1950b. Quantum theory of non-local fields. Part II. Irreducible fields and their interaction'. *Phys. Rev.*, 80 (December): 1047–1052. https://doi.org/10.1103/PhysRev.80.1047.

Advisory Board

Argumenta 10, 1 (2024)
Article Discussion

On Eric Olson's *Parfit's Metaphysics and What Matters in Survival*

The Journal of the Italian Society for Analytic Philosophy

The Fission Argument for the Unimportance of Identity Cannot Be Correct

Harold Noonan

University of Nottingham

Abstract

Eric Olson has made an important addition to the discussion started by Parfit of the argument from the possibility of fission to the unimportance of personal identity. Olson's discussion is challenging. I want, more briefly, to highlight what is the most important consequence of it. This is that it is metaphysically impossible, impossible in the strongest sense, that any version of Parfit's argument from fission can yield his conclusion. Olson argues specifically that this is impossible if what he calls a 'capacious ontology' is assumed. I argue that it is a consequence of Parfit's reasoning that this is so even without the assumption of a capacious ontology.

Keywords: Fission, Identity, What matters, Parfit, Olson.

1. Introduction

Sometimes, occasionally, something new and important is added to a long-running philosophical debate. Eric Olson (2019) has made just such an addition to the discussion started by Parfit (1971) of the argument from the possibility of fission to the unimportance of personal identity.

But Olson's discussion is long, complex and challenging. I want, more briefly, to highlight what is the most important consequence of it—a consequence he does not actually draw out. This is that it is metaphysically impossible, impossible in the strongest sense, that any version of Parfit's argument from fission can yield his conclusion.

The reason for this is that any version of the argument:

(a) has to appeal to the difference between two situations (i) one in which a single brain hemisphere is transplanted (with consequent transfer of psychology) and the other destroyed; and (ii) one in which two hemispheres of a brain are transplanted into distinct skulls (with consequent transfer of psychology)—the fission case, and

Argumenta 10, 1 (2024): 369—373
ISSN 2465-2334

DOI 10.14275/2465-2334/202419.noo
First published: 30 November 2024

(b) must assume (premise 1) that whilst identity is preserved in the first case there is no identity in the second, though (premise 2) everything *that matters* is preserved in the second case as in the first.

Of course, Parfit needs to justify the second premise, that everything in the single hemisphere transplant that matters is preserved in the fission case, as well as the first. He could just insist that only psychological continuity matters. But, as Olson notes, he does not want to do that. I think the best response Parfit has, at this point, is to appeal to our intuition when we think about the possibility *first-personally*: it seems that given a choice between a single hemisphere transplant and fission there is nothing to make it reasonable to choose the former. This seems a good reply (Shoemaker 1984: 119) to a demand for a justification of Parfit's second premise. So, the crux, which Olson is mainly concerned with, is whether the first premise, that identity is not preserved in the fission case though it is in the single hemisphere transplant, can be defended. My claim in what follows is that thinking through Olson's criticism we can see that it cannot be, even if a capacious ontology (as Olson calls it) is not assumed. Note that throughout when I say 'identity' I mean personal identity. A capacious ontologist might say that identity is preserved in fission, but not personal identity. That is, he might say that there is something, one and the same thing, present before the fission and afterwards, but that there is no person present before and after the fission. But that would be implausible, no one does say this and Olson sensibly ignores the possibility.

2. Why the Fission Argument Fails

I now go on to explain all this.

The focus of Olson's argument is, in fact, what he calls "the capacious ontology"—the ontology of a philosopher who thinks that every matter-filled region of space-time contains a material thing which exactly matches its boundaries (Olson 2019: 30). An example of this is the four-dimensional ontology of Lewis (1976) and Quine (1960), in which any shorter-live thing coincident throughout its existence with a longer-lived one is a temporal part of the latter. But Olson uses the term more generally. He makes a convincing case that Parfit accepts the capacious ontology, though without ever arguing for it, but he notes that Parfit is silent on the Lewis–Quine ontology of temporal parts. He also draws attention to Shoemaker (Shoemaker 1984), who also seems committed to a capacious ontology, but vociferously rejects the Lewis–Quine ontology of temporal parts.

Olson then goes on to argue that the defender of the capacious ontology cannot employ Parfit's fission argument to establish the unimportance of identity, in the sense championed by Parfit (so, of course, by assuming the capacious ontology, Parfit has undermined his own argument).

His argument for this claim depends on a careful distinction between what Parfit is arguing for and the (uninteresting) claim he is not arguing for.

Parfit's actual claim Olson expresses as follows:

Strong Unimportance of Identity: What matters in survival is never identity, but only some sort of psychological continuity. Whenever someone has a special prudential reason to care about someone's future, it's not because anyone survives, but only because that future person is psychologically continuous with her.

He distinguishes this from:

> *Weak Unimportance of Identity:* What matters in survival is always identity. Psychological *continuity* is practically important because it secures identity. Whenever someone has a special prudential reason to care about someone's future welfare, it is either because she is the person and thus survives or because someone coincident with her survives. But it is always because someone survives.

According to the weak claim, psychological continuity is not what ultimately matters. What does is identity. But what matters to a person about to fission is not that he, the very same person, exists after the fission. What matters to him is that there is *a* person coincident with him before the fission who exists after the fission, and so persists as one and the same identical thing through the fission. The previously coincident person may or may not be psychologically continuous with himself as he was earlier. This is not important to a person about to fission. What matters to such a person is only that someone coincident with him before exists after. This is not Parfit's claim. It is no one's claim. As Olson puts it, "strong unimportance of identity is a radical challenge to our ordinary thinking about value. The weaker claim is much less interesting. The most likely reaction to it is bafflement. It is unlikely to change our thinking".

But, as Olson explains, Parfit's actual thesis about the unimportance of identity cannot be supported by appeal to the fission argument if the capacious ontology is assumed. According to the capacious ontologist, it is metaphysically necessary that in a case of fission there is survival. So, a thought experiment separating the two factors that might ground what matters—the presence of identity on the one hand (as in the single hemisphere transplant case) and the presence of mere psychological continuity (as in the fission case)—is metaphysically impossible. Granted that nothing is present in the former that matters which is lacking in the latter, we cannot infer that identity is not something that matters since, according to the capacious ontology, there is identity in the latter too.

However, it is obvious that one can think that there is identity in the fission case, i.e., that one can think that someone who exists after the fission in that situation existed before, without endorsing the capacious ontology. One needs not believe that every filled space-time region contains an object which exactly fills it to believe this.

A plausible line of thought that yields the conclusion that if there is someone in the single brain-hemisphere transplant case who survives the transplant then someone who is present after fission in the fission case was there before the fission, goes as follows. First thought. A person cannot go out of existence unless something happens to him. But in the relevant sense something happens to a person only if he undergoes a non-relational change. Nothing thereby happened to Socrates when Theaetetus grew taller than him. Nothing thereby happens to a man when his long-separated wife dies—though he becomes a widower. Nothing happened to the Merry Men when evil Prince John had a sudden change of heart and pardoned them, and the next day, returning to his old ways, reversed the pardon—though the number of outlaws in Sherwood Forest went from 100 to 0 then back up to 100. That a person cannot go out of existence merely because of a relational change is a fact, a necessary fact, about persons. It is not a fact about things generally, it is not, for example, a fact about holes or indentations more

generally.[1] But it is a fact about lots of things other than people: dogs and trees and ships and computers and ashtrays. People are like dogs and trees etc., not like holes. The second thought is simply that if a person (or dog or tree etc.) does not go out of existence at some time in one situation, it cannot go out of existence at that time in any second situation in which nothing happens to it that does not happen to it in the first. This is just part of what it is to be a person or a dog etc.[2]

If this line of thought is accepted, then—even if the capacious ontology is rejected—it must be acknowledged that, in the fission case, there is necessarily someone who exists after the fission who existed before it, if there is a person with such a lifespan in the single hemisphere transplant case. So, we again secure, by Olson's reasoning, that it is impossible for any version of the fission argument to secure Parfit's conclusion, since no thought experiment separating the two factors that might ground what matters is metaphysically possible.

Of course, someone might resist the line of thought just described and insist that a mere relational change can bring a person's existence to an end—persons *are* like holes (he then has to choose whether to say the same of dogs etc., or to accept that persons are unlike dogs). But, apart from a defender of the capacious ontology, who thinks things are constantly going out of existence without any non-relational change happening to them, who would want to say this? This is the line that must be taken by those who endorse a non-branching, no-rival or best candidate, account of personal identity. But those who endorse this are typically capacious ontologists—the most prominent defenders of such an account of personal identity being Parfit himself, and Shoemaker.

I conclude that reflection on Olson's argument should lead to the position that Parfit's fission argument necessarily fails to yield its conclusion. Maybe some other argument will do the job. But Parfit's own additional argument, the argument from below, is much contested, and specifically, as Olson shows, requires the assumption of the capacious ontology and is thus inconsistent with the strong independence of what matters from identity that Parfit believes in. And I know of no other. So, I think that where we are at present is that there is no good reason to accept Parfit's famous claim that identity does not matter.

References

Noonan, H., 2019. *Personal identity*. New York: Routledge.

[1] One can bring a hole into existence by digging. So, a way to cause a hole to cease to exist is to fill it up. But one can also cause it to cease to exist by lowering the ground around it.
[2] This line of thought, of course, derives from Williams (1956–7) and is employed by him in Williams (1970). Noonan (2019: 140) attempts a formulation of the basic principle (as applicable to persons) that can be put as follows in the terminology of this paper: "If two events are parts of the history of a person in one situation they must also be parts of the history of a person in any second situation in which they, and all the events which are part of the history of the person in the first situation, remain present and differ in no non-relational way from the way they are in the first situation". J.R.G. Williams (2013) gives a better formulation which can be put as follows: "If a spatio-temporal region is exactly occupied by a person any duplicate (intrinsically identical) region is exactly occupied by a person or is part of a region exactly occupied by a person".