

G. H. Lewes: Emergentism and The Vitalist Roots of Non-Reductive Physicalism

Alexander Klein, McMaster University

DRAFT. Please do not quote or cite without permission.

25 August 2022

Abstract

The claim that the conscious mind might “emerge” from the physical brain without being reducible to it came in for a revival at the end of the 20th century. As proponents like David Chalmers have long acknowledged, the operative concept of emergence is originally due to figures like C. D. Broad and C. Lloyd Morgan (who are sometimes referred to as “British Emergentists”). But the word “emergence”—and indeed one of the earliest formulations of this concept—is due to the 19th century philosopher and naturalist G. H. Lewes. In this paper I show that Lewes’s concept of emergence is only intelligible against the background of his deeper commitment to vitalism. For Lewes, vitalism is the view that organic matter is governed by a distinct, irreducible set of natural laws, and that these special laws are required to explain how mentality fits into the natural world. I offer an analysis of Lewesian vitalism, and conclude by pointing to some striking similarities between that view and more recent, non-reductive naturalist views about consciousness. More recent advocates of the latter view don't just rely on the concept of emergence in articulating the relationship between consciousness and the underlying brain. They also endorse the strikingly Lewesian, strikingly vitalist view that our inventory of fundamental natural laws must be expanded to include “special” laws that govern the neural phenomena that give rise to conscious mentality.

1. Introduction

“Vitalism” has sometimes been used as an epithet to smear non-reductive, naturalistic views about consciousness (Dennett 1996, Garrett 2006). Arguments purporting to show that consciousness is not reducible to physical facts would also establish that life is not reducible to physical facts, according to critics. But the latter view is vitalism, and vitalism is absurd. In this paper I offer historical considerations that dampen the force of this reductio.

One theoretical innovation embraced by influential non-reductivists is that the conscious mind might “emerge” from the physical brain without being reducible to it. As proponents like

Chalmers have long acknowledged, the operative concept of emergence is originally due to work from the 1920s by figures like C. D. Broad and C. Lloyd Morgan (who are sometimes referred to as “British Emergentists”; see Chalmers 2006). Both Broad and Morgan acknowledged their indebtedness to G. H. Lewes, the 19th-century “positivist” and “empiricist” (his words).

Lewes coined “emergence” as a term of art and was one of the first to develop this philosophical concept in detail (McLaughlin 1992, 58). Lewes appears to have been *the* first to suggest that mentality (more precisely, “sentience”)¹ is a high-level property that emerges from an assemblage of lower-level physiological units (Lewes calls the units “neural tremors”; Lewes 1874, 135 – 36).

That Lewes pioneered the concept of emergence is not (entirely)² controversial. But I shall argue that Lewes’s concept of emergence is only intelligible against the background of his deeper commitment to vitalism.

To see this, note that a cornerstone of Lewesean emergence is that although sentience consists of nothing but a grouping of physiological elements, nevertheless those physiological elements fail to “explain” the emergent effect. We cannot “deduce” the emergent property from what is happening at the base-level (and *mutatis mutandis* for any other “emergent” property with respect to the relevant base-level elements; Lewes 1879, 8).

To make sense of Lewes’s reasons for thinking that emergent properties are in principle not deducible from facts about base-level elements, we must appeal to his vitalism. For Lewes, vitalism is the view that organic matter is governed by a distinct, irreducible set of natural laws, and that these special laws are required to explain how mentality fits into the natural world. Vitalism now has a bad name, suggesting overtones of supernaturalism. But Lewes insisted that vital principles must be natural and not supernatural. Vital laws are not mechanical laws, but vital laws are still natural, he held.

¹ Sentience has both a conscious and a sub-conscious manifestation, for Lewes—see fn. 4, below.

² See fn. XX, below.

Lewes motivated this vitalism by appealing to aspects of animal behavior (especially learning) that fundamentally separate it, he thought, from the behavior of inorganic matter. A capacity to learn makes animal behavior “incalculable,” and this incalculability is characteristic not only of sentient phenomena, but of all vital phenomena (including neurophysiological function) generally. This is why sentient phenomena are impossible to deduce from lower-level facts about the animal “machine,” for Lewes (Lewes 1877b, 324).

Today’s non-reductivists also argue that consciousness must be connected to the physical via unique, fundamental laws of nature. Although Chalmers has his own reasons for thinking special laws are needed to explain the connection between the physiological and the psychological,³ the basic argumentative strategy is remarkably similar to Lewes’s, and this similarity has a historical, causal explanation.

So how does this historical connection help defuse the *reductio-ad-vitalism* to which non-reductivists are subject today? I agree that non-reductivists (at least those committed to emergence) can fairly be saddled with a form of vitalism—with Lewes’s specific form of vitalism, which clearly informed the development of the very idea that mind “emerges” from neural events. Now, it might or might not be true that special, distinctive laws are needed to explain the connection between minds and bodies (I remain agnostic on this question). But I shall argue that if one already finds this idea in Chalmers to be absurd or supernatural, the move to associate this position with “vitalism” adds little more than a bogey word. And if one is not inclined regard as supernatural the non-reductivist’s claim that special psychophysical laws must connect brain and consciousness, then one should not tremor at the suggestion that this view is linked—both historically and conceptually—with vitalism.

2. Lewes on Emergence

³ Chalmers anticipates the *reductio-ad-vitalism* at (Chalmers 1995, 208 – 09), and later offered a brief response to Dennett’s version of the *reductio* at (Chalmers 2010, 29 – 30).

By at least 1874, when the first volume of the first series of his *Problems of Life and Mind* appeared, Lewes was describing consciousness and sentience as things that “emerge” from organic processes. He wrote:

The great problem of Psychology as a section of Biology is ... to develop all the psychical phenomena from one fundamental process in one vital tissue. The tissue is the nervous: the process is a Grouping of neural units. A neural unit is a tremor. Several units are grouped into a higher unity, or neural process, which is a fusion of tremors, as a sound is a fusion of aerial pulses; and each process may in turn be grouped with others, and thus, from this grouping of groups, all the varieties *emerge*. What on the physiological side is simply a neural process, is on the psychological side a sentient process. We may liken Sentience to Combustion, and then the neural units will stand for the oscillating molecules. Sentience may manifest itself under the form of Consciousness, or under that of Sub-Consciousness which may be compared to Combustion manifesting itself in Flame and in Heat. (Lewes 1874, 135 – 36, italics mine)

Here we have Sentience—which can “manifest” in either a conscious or sub-conscious fashion⁴—“emerging” from neural tremors.⁵ Emergence is portrayed as a relationship between a “higher unity,” or in other words a “Grouping,” and the elements that compose it, where properties that can be predicated of the whole group cannot be predicated of the simpler component elements by themselves. Lewes gives an example of another emergent phenomenon—combustion emerging from underlying “oscillating molecules.” None of the oscillating molecules by themselves constitute combustion. Combustion is a phenomenon that

⁴ Lewes offered an extended discussion of his distinction between conscious and sub-conscious sentience in a *Mind* article (Lewes 1877a) which was excerpted from his then-forthcoming (Lewes 1877b).

⁵ William James had no patience for Lewes’s distinction between conscious and sub-conscious sentience. In a review of (Lewes 1877b), James wrote: “the vagueness of that ‘sentience’ which, as distinguished from ‘consciousness,’ Mr. Lewes affirms to be the property of all nerve tissue (for so we now understand him), and the vagueness of the distinction between it and consciousness as he elaborates it, are fairly lamentable. What, indeed, is gained by merely saying that when a man’s dorsal marrow is crushed his legs still feel, though he does not? What we want to know is how they feel, and who they are when they feel, and how it comes that only when they are in organic continuity with the brain their feelings form part of an ego” (ECR 1877, 344). James’s rejection of sub-conscious sentience is not surprising in light of his own attacks on unconscious mentality (see Klein 2020).

can be attributed, under some unspecified conditions, to groups of oscillating molecules (he thinks)—that is, combustion emerges only at a macro level. Similarly, although nervous “tremors” are not themselves sentient, sentience can be attributed, again under some unspecified conditions, to groups (or “groupings of groups”) of nervous tremors. It, too, emerges only at a macro level. For Lewes, that macro-level is the entire nervous system, and this is one reason he insists on spinal sentience (Klein 2018)—he denies that sentience is localized in the brain (though he does think the brain plays a leading role in organizing sentient response, as we will see).

Today we might simply say that heat “supervenes” on oscillating molecules—it is not possible for there to be a change in heat without there being a change at the microphysical level.⁶ Because of this supervenience relation, it is often said that heat is *weakly* emergent from molecular motion, and Lewes thinks sentience and consciousness are at least weakly emergent (to use today’s distinction)⁷ from neural tremors. He writes:

I take for granted that the living nervous mechanism has one general mode of action which may be called Sensibility. This general mode manifests itself in sensible tremors, groups of such, and groups of groups—in sensations, perceptions, emotions, conceptions—which are never manifested apart from this mechanism, and *which vary with every variation in the molecular movement of that mechanism.* (Lewes 1875, 459 – 60, italics mine)

⁶ Care must be taken with the term “supervenience,” which was coined by the British Emergentist C. Lloyd Morgan, and originally used simply as a synonym of “emergence”—Lloyd Moran apparently used the phrase to indicate properties that are higher-order groupings of base properties, but that *cannot* be reduced to those base properties (McLaughlin 1992). I am using the term in its different, contemporary understanding, to mean a relation between two properties such that one (the supervening property) cannot change without the other (the base property) changing. That is not Morgan’s sense, since supervening properties in our contemporary sense can at least arguably be reduced to base properties.

⁷ Strong emergence is the case where a high-level phenomenon supervenes on a lower-level phenomenon, but the higher-level phenomenon is not deducible even in principle from the lower-level facts and the laws governing that level. Weak emergence is the case where supervenience obtains, and the higher-level phenomena may be in some way surprising or unexpected, but is nevertheless deducible in principle from the lower-level phenomena and the laws governing that level. See (Chalmers 2006).

We can say that the form of supervenience Lewes suggests is “strict” and not just “strong” in our contemporary sense, especially in light of the italicized portion of the quotation. Lewes held that every variation in sensation must (nomologically) go along with a variation in molecular movement in the organism—but more surprisingly, he insisted that every variation in molecular movement *also* must make a difference in sensation, too. We will see below that this strict, bi-directional form of supervenience will create philosophical trouble for him.

Supervenience alone—even Lewes’s strict version of this relationship—does not yet get us the distinctive notion of emergence—*strong* emergence, as it is often called today. Strong emergence would require not just supervenience, but also the denial of reduction. And Lewes indeed did deny that sentience can be fully *reduced* to molecular motion, in the following sense.⁸ Lewes suggested that even though there can be no change in sentience without a change in neural tremors (and vice versa), he expressly denied that the state of the neural tremors *explains* the sentient states—we cannot deduce one from the other. Thus in his posthumous work *Problems of Life and Mind, Third Series*, he contended that even though “spiritual facts” arise in virtue of lower-level biological processes, those spiritual facts “can never be explained by, or reduced to, material facts.” He went on:

I ... admit that no deductions from what is known objectively of the material mechanism will explain the phenomena of sensibility, as states of consciousness, any more than anatomical knowledge of an organ alone will enable us to deduce its function. (Lewes 1879, 8)

That Lewes both embraced supervenience *and* denied reducibility is why he should be regarded as a founder of modern emergentism, which enjoyed a vogue in the opening decades of the 20th century among biologists and philosophers like C. Lloyd Morgan and C. D. Broad, and which has lately been revived.⁹

⁸ For two contemporary accounts of *emergence* that analyze that relation as supervenience plus non-reduction, see (Chalmers 2006, Kim 2006). Supervenience plus non-reduction is what Chalmers calls “strong emergence.”

⁹ For various acknowledgements of Lewes’s importance to this tradition, see e.g. (Clayton and Davies 2006, Feinberg and Mallatt 2020, McLaughlin 1992).

Morgan credited Lewes with coining the term “emergent” (for discussion see McLaughlin 1992, 58). Indeed, Lewes argued that we must distinguish two distinct kinds of causal products, “resultants and emergents” (Lewes 1875, prob. 5 ch. 3). Lewes used “resultant” to mean an effect that is “homogeneous and commensurable” with its cause (Lewes 1875, 413), and his central example was that of a vector sum (today the word “resultant” is still used in physics to describe the product of two forces applied to a body at once). He wrote:

the resultant motion of two impressed forces is the diagonal of those directions which the body *would* take were each force separately applied. Every resultant is either a sum or a difference of the cooperant forces; their sum, when their directions are the same—their difference, when their directions are contrary. Further, every resultant is clearly traceable in its components (Lewes 1875, 413)

A vector sum is a *resultant* in that we can precisely and clearly trace how the effect is brought about by each causal ingredient.

But not every effect is a resultant in the above sense:

... although each effect is the resultant of its components, the product of its factors, we cannot always trace the steps of the process, so as to see in the product the mode of operation of each factor. In this latter case, I propose to call the effect an emergent. (Lewes 1875, 412)

The classic example of an emergent—both for Lewes and for Mill, to whom Lewes was indebted (McLaughlin 1992)—was chemical composition. Lewes gave the example of gold being turned into a chloride when submerged in a bath of hydrochloric and nitric acid (a bath of *aqua regia*—literally royal water, as it is known due to its surprising capacity to dissolve otherwise inert, “noble” metals).¹⁰ “Who could foresee” this effect “before experiment,” Lewes asked rhetorically (Lewes 1875, 414).

¹⁰ When hydrochloric and nitric acid are combined, the chloride produced (*aqua regia*) is chloroauric acid, HAuCl_4 .

These passages illustrate that Lewes often gave an epistemic gloss to this distinction—emergents are those effects linked to causes in a way we do not fully *understand*. Hence he complains about the illicit use of words like “effect” to “fill up this gap in our knowledge” when we look at cases like chemical composition. He prefers the word “emergent” to “effect” when there is a gap in our knowledge. We may in the future come to take some cases of chemical composition to be resultants and not emergents as our knowledge improves, Lewes held, which again suggests an epistemic understanding of the distinction (Lewes 1875, 414).¹¹

Still, in other places Lewes gave an *ontological* gloss to his distinction between resultants and emergents. Vector sums involve “adding measurable motion to measurable motion,” and Lewes emphasized that the effect flows from a combination of causes that are *alike in kind*. He held this to be typical of—perhaps essential to—resultant causation. In contrast, he portrayed emergents as essentially involving “a cooperation of things of unlike kinds,” giving this example: “Unlike as water is to oxygen or hydrogen separately, or to both when uncombined, nothing can be more like water than their combination, which is water” (Lewes 1875, 413 – 14).¹²

Perhaps Lewes’s epistemic way of cashing out the relationship speaks to his views about how we *detect* that a causal relationship might be resultant or emergent, but that the ontological treatment gives his account of what *makes* a causal product either a resultant or an emergent. On

¹¹ McLaughlin suggests that British emergentism died with the rise of quantum mechanical explanations of chemical bonding, and emphasizes that “no even remotely plausible micro-explanations of chemical bonding” existed in Lewes’s day (McLaughlin 1992, 54 – 55). In other words, once we *know* how chemical bonding works, chemical composition ceases to be unforeseeable. He may be right that as a historical matter, the rise of a good explanation for a favorite example of purported emergence (chemical bonding) made the view less attractive. But I note that at least when Lewes uses “emergence” in an epistemic sense, he explicitly says that chemical bonding may well count as an emergent phenomenon today but will count as a resultant as our knowledge grows. So McLaughlin’s point about chemical bonding seems plausible from a historical standpoint, but it does not threaten Lewes’s philosophical position. One can give up chemical bonding as an illustrative example of emergence while still maintaining that other phenomena like consciousness do emerge from the physical.

¹² This usage appears to be indebted to Mill’s notion of a “heteropathic effect.” As (McLaughlin 1992, 58 – 59) points out, Mill had developed a similar idea in book three chapter 6 of his 1843 *Logic* (“On the Composition of Causes”). Indeed, the use of vector sums to illustrate what Lewes calls “resultants” comes straight from Mill (Mill 1843/1974, 370 – 71). Still, McLaughlin overstates his case when he says that Lewes’s “notion of an emergent is ... just Mill’s notion of a heteropathic effect” (McLaughlin 1992, 58). While Mill does suggest that life and mentality may be governed by heteropathic laws, he does not (so far as I can find) analyze mind or sentience as a higher-level property emerging from an *assemblage* of unlike elements (like neural tremors, on Lewes’s telling). This is Lewes’s important, distinctive contribution to the discussion.

this way of reading Lewes, if we do learn how to perfectly foresee the result of combining hydrochloric and nitric acid, *what* we learn is that we were mistaken in *regarding* this combination as a case of emergence.

In any case, Lewes clearly sees the relationship between sentience and neural tremors as one of emergence. This is because a) a strict form of supervenience obtains, and b) we cannot “deduce” (Lewes’s word) facts about sentience from facts about nerve-firings.

Now it must be said that the version of emergence Lewes proposes is more counter-intuitive than emergentists today (who sometimes cite Lewes approvingly, if in passing) might realize. For Lewes’s view entails that emergence must be bi-directional. Given how strictly mental events are connected with neural events, his account would have to count neural events as “emergent” from mental events just as much as it would count mental events as “emergent” from neural. That is an odd consequence, and it stems from his excessively strict conception of supervenience.

Here is a way to think about the problem. The usual cases of purported supervenience that we care about are uni-directional—heat supervenes on molecular motion, but molecular motion does not in turn supervene on heat. Supervenience obtains because a change in heat requires a change in molecular motion, but not vice versa. Identity can be regarded as a special case of supervenience that is bi-directional. If the person the name “Scott” denotes is identical with the person the phrase “The Author of *Waverly*” denotes, then (setting aside intensional contexts) any change in Scott requires a change in the Author of *Waverly*—if Scott cuts his hair, then so does the Author. So Scott supervenes on the Author. But in this case, the supervenience relationship goes both ways—if the Author cuts *his* hair, then so does Scott. So the Author also supervenes on Scott.

The relationship Lewes is proposing between sentience and groups of nervous tremors—what I am calling strict supervenience—is bi-directional, like the identity relationship, even though it is ultimately a weaker relationship than identity. It is bi-directional for the simple

reason that Lewes insists that a change in sensation always goes with a change in neural firing, *and vice versa*—every change in neural firing goes with a change in sensation. Recall the passage I quoted from (Lewes 1875, 459 – 60).

Lewes is not offering an identity theory, though. He sees sentient states and neural tremors as two different “aspects” of one process (Lewes 1875, 483, 1877b, 390, 403). Hollowness and brownness are two aspects of my violin; but obviously hollowness is not identical with brownness.¹³

So we seem to have a relationship of supervenience that is bi-directional—the sensational aspect supervenes on the neural, *and* the neural aspect supervenes on the sensational. What is more, if we cannot deduce the mental state from the neural state, presumably we also cannot deduce the neural state from the mental. Hence he must accept that the neural emerges from the mental every bit as much as the mental emerges from the neural. I do not believe that was his intention.¹⁴

Some might suspect that something has gone wrong in my reading. Lewes initially presents resultants and emergents as two different kinds of *causation*, and causation is typically thought to be uni-directional.¹⁵ If x causes y, then y does not cause x. So perhaps we should read Lewes as silently stipulating that y does not emerge from x unless x causes y. We would then have a three-part definition of Lewesean emergence, where y emerges from x if and only if a) y strictly supervenes on x, b) y cannot be deduced from x, and c) x causes y.

The trouble is that Lewes eventually rejected stipulation c) quite explicitly:

¹³ Lewes’s dual-aspect account of sentience and neural tremors must be independent of his supervenience claim. Painting my violin requires no change in its hollowness; and lopping off the bottom of the violin so it is no longer hollow would not require any change in color, obviously.

¹⁴ Today, emergentists employ more plausible and sophisticated accounts of supervenience, and this marks an important step forward from the version of emergentism Lewes developed. As Chalmers has pointed out to me, the directionality problem does not arise if one thinks that the same mental state can emerge from *different* brain states (personal communication). For an emergentist who accepts this version of mind-brain supervenience—and I take it that includes most emergentists today—brain states would nomologically necessitate the mental states, but the opposite would not be true.

¹⁵ There are philosophers who claim that there are genuine cases of mutual causation; for a discussion and criticism of cases of this purported phenomenon, see (Frankel 1986).

It is ... indifferent whether we say a sensation is a neural process, or a mental process: a molecular change in the nervous system, or a change in Feeling. It is either, and it is both, as I have elsewhere explained.¹⁶ There, it was argued that the current hypothesis of a neural process *causing* the mental process—molecular movement being in some mysterious way *transformed* into sensation—is not only inconceivable, but altogether unnecessary; whereas the hypothesis that the two aspects of the one phenomenon are simply two different expressions, now in terms of Matter and Motion, and now in terms of Consciousness, is in harmony with all the inductive evidence. (Lewes 1877b, 403)

Lewes is expressing a familiar worry, that it is “inconceivable” that a mental event could causally alter a physical event, or that a physical event could causally alter a mental event. Lewes’s dual-aspect view is apparently designed to skirt such problems. An attractive feature of standard dual-aspect views of minds and bodies is that they do not require mind-body or body-mind causation. But this means that when Lewes characterizes minds *emerging* from neural tremors, there is really no causation involved at all, and he cannot accept (c), above.

Indeed, while Lewes had initially presented his distinction between resultants and emergents as two different varieties of causation, he went on to kick out the ladder in the same volume, arguing that *no* instance of causation should be viewed as a transformation between two distinct events. Instead, he contended that causation is always “one thing differently viewed” (Lewes 1875, 460). So here we find another surprising doctrine in Lewes—causation itself is not a directional process at all, but rather a kind of redescription. So even if we try to ignore his explicit denial (at Lewes 1877b, 403) that there is a causal relationship between the mental and the neural, and try to help him gain back directional emergence by introducing a causal requirement on his behalf, even that will not work, for his own account of causation also seems to be bi-directional.

¹⁶ Here Lewes refers the reader to (Lewes 1875, 457).

Thus we cannot preserve directional emergence on Lewes's behalf by adding in a causation requirement. But this leaves Lewes with an account that is committed to emergence being a reciprocal relationship, at least when it comes to the neural and the mental.

3. Lewes on Vitalism

Lewes called himself a “vitalist” who distinguished two kinds of natural “principles” that govern the function of inorganic and organic matter, respectively. Neural tremors and other physiological functions are not *physical* phenomena, in his view, but *vital* phenomena that obey a distinctive set of natural laws. These laws only apply to organic matter. It is important to see that not only sentience, but the neural base from which Lewes thinks sentience emerges, is supposed to be vital and not physical.

Vitalism now has a bad name, harboring overtones of supernaturalism. But Lewes insisted that vital principles must be “intra-organic” and not “extra-organic” (Lewes 1877b, 324)—that is, they must be natural and not supernatural. Lewes rejected a strong form of Vitalism (which he typically referred to with a capital “V”) that portrayed the laws governing organic matter as somehow *outside* of or *above* nature.¹⁷ Vital laws are not mechanical laws, but vital laws are still natural, he held.

The so-called automatism controversy of the 1870s pitted vitalists like Lewes against mechanists like T. H. Huxley (Klein 2018). But it's important to see that this was a debate between naturalists. *Both* sides denied that mental phenomena involve any supernatural substance or principle. The question was how to understand mind's place *in* nature, not whether it somehow intervenes from above. Lewes and allied vitalists insisted that there are two fundamentally different kinds of laws of nature—mechanical and vital laws. Mechanists only accepted the existence of mechanical laws.

¹⁷ A helpful discussion of Lewes's vitalism can be found in (Price 2014, 112 – 14).

The distinctive character of vital laws helps explain why facts about sentient states cannot be “deduced” from facts about vital states (in particular, from facts about the state of a group of neural tremors—see above), on Lewes’s view. He accepted that organisms are “machines” in that “the adjustment of the parts is a condition of mechanical action.” But vital machines are different from automata:

One cardinal difference is that the combination of the parts is in the [inorganic] machine a fixed, in the organism a fluctuating adjustment; and this fluctuation is due to certain vital processes subjectively known as *sensitive guidance*. Hence machines have fixed and calculated mechanisms; whereas organisms are variable and to a great extent incalculable mechanisms. (Lewes 1877b, 324)

The key difference between inorganic and organic mechanisms is that only the latter are capable of *learning* (Lewes 1877b, 326), and learning is what introduces this “incalculable” component to an organism’s reaction to stimuli.

Incalculability is not mere randomness. What Lewes had in mind is incalculability in the sense that the same stimulus may elicit two different stimuli in an untrained and a trained animal. He offered a simple illustration: “We can conceive an automaton dog that would bark at the presence of a beggar; but not of an automaton dog that would bark one day at the beggar and the next day wag his tail, remembering the food and patting that beggar had bestowed” (Lewes 1877b, 386). For Lewes, the real dog is not like an automaton that has an extra, vital process—a sensational, non-physical process—somehow operating alongside material processes. Instead, Lewes thought that organic and inorganic matter were subject to different sets of natural laws entirely.

Mechanical laws are such that if we know the laws and we know the initial conditions, we can deduce what happens next. If we know enough physics and we know how the automaton dog is constructed, then we can precisely and accurately predict how the automaton dog will react to a given stimulus.

But vital laws are not like this. To Lewes, the fact that organisms can “adapt” or “learn” to cope with the same input in different ways, as the real dog reacts to the beggar on subsequent occasions, suggests that organic matter *on the whole* is subject to a different set of natural laws than an inorganic automaton. And one key, distinguishing mark of the vital laws that govern organic matter, on Lewes’s view, is this kind of unpredictability.

4. Reflection on Non-Reductive Physicalism

Lewes’s position may finally seem interesting but outdated, since all the world has given up on special “vital” principles. But strong emergence has defenders in high philosophical places. Its proponents don’t identify as vitalists, but they do argue that special natural laws—non-physical, natural laws—are needed to explain the systematic connections between consciousness and physical bodies.

Here is Chalmers’ way of articulating the idea:

[C]onsciousness ... supervenes on the physical domain. But importantly, this supervenience holds only with the strength of laws of nature (in the philosophical jargon, it is natural or nomological supervenience). In our world, it seems to be a matter of law that duplicating physical states will duplicate consciousness; but in other worlds with divergent laws, a system physically identical to me might have no consciousness at all. This suggests that the lawful connection between physical processes and consciousness is not itself derivable from the laws of physics but is instead *a further basic law or laws of its own*. The laws that express the connection between physical processes and consciousness are what we might call fundamental psychophysical laws.

I think this account provides a good general model for strong emergence. We can think of strongly emergent phenomena as being systematically determined by low-level facts without being deducible from those facts. In philosophical language, they are naturally but not logically supervenient on low-level facts. In any case like this,

fundamental physical laws need to be supplemented with further fundamental laws to ground the connection between low-level properties and high-level properties. (Chalmers 2006, 247, *my italics*)

Chalmers is alluding to his own argument that (in a nutshell) unconscious zombies who behave exactly like us are conceivable, and therefore possible; and so there must be some special natural laws in our world that aren't physical laws, and that ground systematic connections between physical states and conscious states.¹⁸ An affinity with vitalism of the sort Lewes defended is evident.¹⁹

Chalmers thinks consciousness would emerge from any two systems that are functionally similar in a "fine-grained" way, so that in principle a silicon chip that functionally duplicated a brain's neural structure would give rise to the same experiences (Chalmers 2007, 364). So the special laws he thinks are needed to explain consciousness aren't tied to organic matter only, as they are for Lewes.

But other non-reductivists today, like Searle, do think the (purported) special laws that would explain consciousness must be tied to a specific kind of matter, which would include "brain stuff" in particular (Searle 1984, 40, 2007). Searle calls his position "naturalistic dualism," and I take his approach to be even more in the spirit of Lewes.

So who cares?

¹⁸ For a more general schema for arguments like this in which zombies play a leading role, see (Bailey 2006, 484). Incidentally, we might be tempted to regard pithed frogs as really-existent zombies. But the frogs played quite a different role in the 19th-century context compared to the role they more typically play today, namely as figuring into arguments for non-reductive views of consciousness. For it was Huxley and not Lewes who would have been happy to regard vivisection as creating zombie frogs. Lewes thought pithed frogs illustrate how *unlikely* it is that physiological function could remain intact without the concurrent existence of some form of sentience.

¹⁹ Chalmers explicitly rejects any association with vitalism, claiming the latter movement only sought special laws to explain "the many remarkable functions associated with life, such as complex adaptive behaviour and reproduction" (Chalmers 1995, 209). He distinguishes these functions from phenomenal consciousness, arguing that the former but not the latter can be explained simply by identifying the underlying mechanisms that perform the relevant function (Chalmers 1995, 202). But this attempt to dissociate from vitalism is potentially misleading, since it relies on a distinction (between what we now call access and phenomenal consciousness) that Lewes himself did not draw. Lewes sought to explain adaptive behavior along with phenomenal consciousness, together, by appeal to special vital laws. Chalmers' position is different in appealing to special laws only to explain phenomenal consciousness. But that is to say that Chalmers offers a narrower, more refined version of Lewes's basic explanatory strategy.

As I've mentioned, "vitalism" has long been used as an epithet to smear non-reductive physicalists (Dennett 1996, Garrett 2006). This argument trades on a narrow understanding of "vitalism" as entailing a commitment to Driesch-style entelechies.

In fact, the British emergentists who developed Lewes's ideas were contemporaries of Driesch (1867 – 1941). And they positioned their Lewes-inspired views as a *rejection* of Driesch's entelechies.

For instance, C. D. Broad attacks what he calls "substantial vitalism"—a view he pins on Driesch. Substantial vitalists claim that to explain "the behavior of living bodies" we must posit "the presence in them of a peculiar component, often called an 'Entelechy', which does not occur in inorganic matter or in bodies which were formerly alive but have now died" (Broad 1925, 56). Instead Broad calls his own alternative view "Emergent vitalism," giving a nod to C. Lloyd Morgan and Samuel Alexander (58). Both Alexander and Lloyd Morgan say their emergentism is indebted to Lewes (McLaughlin 58).

You may doubt that today's non-reductivists have a serious intellectual debt to the vitalism of Lewes, Broad, Morgan, and Alexander. Some of today's cutting-edge non-reductivists (including Chalmers) put the concept of emergence to work in the context of a sophisticated position they call "Russellian Monism." Surely they owe a more substantial debt to Russell's (post 1919) philosophy of mind than to these relatively more obscure, professed vitalists.

In response, let me quote from Russell's 1921 *Analysis of Mind* (the first substantial attempt to articulate his neutral monism). Here Russell is giving a behaviourist-style account of a *goal* or *desire*:

A mental occurrence of any kind of sensation, image, belief, or emotion may be a cause of a series of actions, continuing, unless interrupted, until some more or less definite state of affairs is realized. Such a series of actions we call a "behaviour-cycle." ... *The actions constituting the cycle must not be purely mechanical, i.e. they must be bodily movements*

in whose causation the special properties of nervous tissue are involved. (Russell 1921, 75, my italics)

What are these special properties instantiated in nervous tissue? They are *vital* properties.

We may ... distinguish “vital” from mechanical movements by the fact that vital movements depend for their causation upon the special properties of the nervous system, while mechanical movements depend only upon the properties which animal bodies share with matter in general. (Russell 1921, 47)

In fact, the distinction between mechanical and vital laws is absolutely fundamental to the aim of *Analysis of Mind*, which is to explain the relationship between psychology and physics (Russell 1921, 5 – 6)—or, as he also puts it, he wants to find out what it is “that characterizes mind as opposed to matter?” His final answer is both pithy and vitalist: “psychology and physics are distinguished by the nature of their causal laws, not by their subject matter” (Russell 1921, 287).

It should be clear that even well into the twentieth century, Lewes-style vitalism wasn’t an uncommon position, even among committed naturalists like Russell. More to the point, vitalist thinking was conceptually ineliminable from pioneering forms of both classic emergentism and even related, more sophisticated forms of Russellian Monism.

So where does this leave us with respect to the *reductio-ad-vitalism* leveled at today’s non-reductivists? The *reductio* asserts that the same reasoning Chalmers thinks establishes non-reductivism also would establish vitalism, and vitalism is (purportedly) absurd. But vitalism isn’t just a logical *consequence* of the view that special laws are needed to explain the mind’s relation to the brain. The postulation of special mental laws *already is vitalism*. And it’s exactly the kind of vitalism—naturalist vitalism—that this tradition has been committed to from the start.

Today I haven’t defended either non-reductive naturalism or Lewes’s emergent vitalism. My point is that the attempted *reductio* is effectively empty. To attach the label “vitalism” to the

view that non-physical laws are needed to explain the mind is simply to bring back a name that happens to have fallen out of fashion.

- Bailey, Andrew. 2006. "Zombies, Epiphenomenalism, and Physicalist Theories of Consciousness." *Canadian Journal of Philosophy* 36 (4):481-509.
- Chalmers, David. 1995. "Facing up to the Problem of Consciousness." *Journal of Consciousness Studies* 2 (3):200-19.
- . 2006. "Strong and Weak Emergence." In *The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion*, edited by Philip Clayton and Paul Davies, 244-54. Oxford: Oxford University Press.
- . 2010. *The Character of Consciousness*. Oxford: Oxford University Press.
- Clayton, Philip, and Paul Davies. 2006. *The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion*. Oxford: Oxford University Press.
- Dennett, Daniel. 1996. "Facing Backwards on the Problem of Consciousness." *Journal of Consciousness Studies* 3 (1):4-6.
- Feinberg, Todd E., and Jon Mallatt. 2020. "Phenomenal Consciousness and Emergence: Eliminating the Explanatory Gap." *Frontiers in Psychology* 11 (1041):1-15.
- Frankel, Lois. 1986. "Mutual Causation, Simultaneity and Event Description." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 49 (3):361-72.
- Garrett, Brian Jonathan. 2006. "What the History of Vitalism Teaches Us About Consciousness and the 'Hard Problem'." *Philosophy and Phenomenological Research* 72 (3):576-88.
- Kim, Jaegwon. 2006. "Being Realistic About Emergence." In *The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion*, edited by Philip Clayton and Paul Davies, 189-202. Oxford: Oxford University Press.
- Klein, Alexander. 2018. "The Curious Case of the Decapitated Frog: On Experiment and Philosophy." *British Journal for the History of Philosophy* 26 (5):890-917.
- . 2020. "The Death of Consciousness? James's Case against Psychological Unobservables." *Journal of the History of Philosophy* 58 (2):293-323.
- Lewes, George Henry. 1874. *Problems of Life and Mind, First Series: The Foundations of a Creed*. 2 vols. Vol. 1. London: Trübner & co.
- . 1875. *Problems of Life and Mind, First Series: The Foundations of a Creed*. 2 vols. Vol. 2. London: Trübner & co.
- . 1877a. "Consciousness and Unconsciousness." *Mind* 2 (6):156-67.
- . 1877b. *Problems of Life and Mind, Second Series: The Physical Basis of Mind*. London: Trübner & Co.
- . 1879. *Problems of Life and Mind, Third Series*. 2 vols. Vol. 1, problem the first: the study of psychology. London: Trübner & Co.
- McLaughlin, Brian P. 1992. "The Rise and Fall of British Emergentism." In *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*, edited by Ansgar Beckermann, Hans Flohr and Jaegwon Kim, 49-93. Berlin: De Gruyter.
- Mill, John Stuart. 1843/1974. *A System of Logic Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*. Edited by J. M. Robson. 33 vols. Vol. 7, *Collected Works of John Stuart Mill*. Toronto: University of Toronto Press.
- Price, Elfed Huw. 2014. "George Henry Lewes (1817-1878): Embodied Cognition, Vitalism, and the Evolution of Symbolic Perception." In *Brain, Mind and Consciousness in the History*

Klein / DRAFT / *do not quote or cite without permission*

of Neuroscience, edited by C. U. M. Smith and Harry Whitaker, 105-23. Dordrecht:
Springer.
Russell, Bertrand. 1921. *The Analysis of Mind*. London: Allen & Unwin.