The Conceptual Span and Plausibility of Emergence

Applied to the Problem of Mental Causation

by

Thomas Fournier

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Approved April 2013 by the
Graduate Supervisory Committee:

Bernard W. Kobes, Chair
Steven L. Reynolds
Brad Armendt

ARIZONA STATE UNIVERSITY

May 2013

ABSTRACT

This thesis explores the conceptual span and plausibility of emergence and its applicability to the problem of mental causation. The early parts of the project explicate a distinction between weak and strong emergence as described by Jaegwon Kim. They also consider Kim's objections regarding the conceptual incoherence of strong emergence and the otiose nature of weak emergence. The paper then explores Mark Bedau's in-between conception of emergence and ultimately finds that middle conception to be both coherent and useful. With these three emergence distinctions in hand, the thesis goes on to explore Evan Thompson's recent work – *Mind in Life* (2010). In that work, Thompson advances a strong emergence approach to mind, whereby he concludes the incipient stages of cognition are found at the most basic levels of life, namely – biologic cells. Along the way, Thompson embraces holism and a nonfundamental/nonhierarchical physics in order to counter Jaegwon Kim's objections to the notion of downward causation needed for strong emergence. The thesis presents arguments against Thompson's holism and nonfundamental physics, while supporting his assertion regarding the incipient stages of cognition. It then combines an important distinction between mental causation and the experience of mental causation with Thompson's notion of incipient cognition to arrive at a dual realms approach to understanding mental causation.

ACKNOWLEDGMENTS

I am indebted to my committee members, Bernard Kobes, Steven Reynolds, and Brad Armendt, for providing helpful criticisms and suggestions resulting in improvements to this work, and for pointing me to additional research resources. I owe special thanks to Bernard Kobes for inspiring my interest in the topic of emergence and, more generally, philosophy of mind. His lectures, discussions and feedback on my prior work have been an invaluable contribution to my academic development, and his respectful demeanor in rebuttal to my sometimes naïve assertions has inspired my continuing interest in the study of philosophy as a means to personal growth.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

PREFACE

This paper comprises five chapters exploring the concept of emergence and its potential applicability to the problem of mental causation. Chapter 1 explores the conceptual span of emergence focusing primarily on the work of Jaegwon Kim and Mark Bedau. Chapter 2 explores Evan Thompson's notion of emergent embodied mind and Kim's worry over downward causation. Chapter 3 considers and rebuts Thompson's attempts to justify downward causation. Chapter 4 considers what does and does not work regarding Thompson's embodied mind theory and offers an approach to mental causation that employs some of Thompson's work with certain observations about causation. Chapter 5 summarizes the work in prior chapters.

Thesis 1: Jaegwon Kim explores emergence via a bipartite/weak-strong distinction. In doing so he misses an important middle ground offered by Mark Bedau's tripartite view. Without this middle ground, Kim too hastily relegates all claims of emergence into one of two bins, namely: incoherent metaphysics or unremarkable platitudes about our epistemic limitations. I will argue that Kim's claim of incoherence does not stand against close examination in regard to at least two conceptions of strong emergence. I will also argue that a more careful consideration of Bedau's computationally irreducible emergence is both coherent and genuinely worthy of remark, for it illuminates an objective ontology of macro-level structures with irreducible overarching identities.

Thesis 2: Evan Thompson supports a version of strong emergence via dynamic systems complexity resulting in new causal powers through globally emergent dynamic structures. He argues living organisms are dynamic, complex systems displaying reflexive co-emergence of self and cognitive world. On this view, biological cells possess the incipient stage of cognition. Emergent global process structures lead Thompson to posit holism as a feature of nature, and nature as a web of processes *without* fundamental particulars at base. From that platform, he argues against Jaegwon Kim's criticisms of strong emergence, downward causation, and overdetermination.

I will argue in defense of Kim's downward causation worry and against Thompson's notions of holism and nonfundamental physics. Despite my stand against Thompson's physics, I will argue in favor of his views about the incipient stages of cognition. By combining that portion of his views with my own observation about the important distinction between mental causation and the *experience* of mental causation, I will suggest a dual-realms approach to mental causation. This approach offers a more plausible notion of embodied mental causation by not requiring downward causation of the productive/generative kind from the nonphysical realm of conscious awareness to the concrete realm of physical entities like neurons and muscles.

CHAPTER 1

NON-PLATITUDINOUS COHERENT EMERGENCE IN BETWEEN WEAK AND STRONG

1.1. Overview of Kim's Claim of Incoherence:

In ""Supervenient and Yet Not Deducible": Is There a Coherent Concept of Ontological Emergence?" Jaegwon Kim casts doubt on the coherence of particular emergence views that hold certain types of higher level phenomena supervene on lower level phenomena while not being reducible to them.[1] He lays groundwork by first classifying emergence into categories of 'epistemological/weak' versus 'metaphysical/strong'. These distinctions are as follows.

Weak/epistemological emergence involves ignorance and/or intrinsic limitations with respect to our epistemic abilities such that higher level phenomena seem to be novel or unpredictable given information about their basal origins. Kim writes, "The main point then is that we, as cognizers, cannot get there from here—that is, get to higher-level emergent phenomena from information about the lower-level base phenomena."[2] Kim expresses considerable apathy toward weak emergence claims – calling them unremarkable, platitudinous and philosophically unexciting.[3]

In contrast, strong/metaphysical emergence involves *objective* newness and yet irreducibility with respect to certain higher level phenomena emerging from the lower

[1] Jaegwon Kim, "“Supervenient and Yet Not Deducible”: Is There a Coherent Concept of Ontological Emergence?” *Essays in the Metaphysics of Mind* (Oxford: Oxford University Press, 2010) 85-104.

[2] Kim, 87.

[3] Kim, 104.

level phenomena. Despite this irreducibility, the higher level phenomena are dependent on their lower level bases. Furthermore, the irreducibility is not just a human epistemological limitation. It is metaphysical. Therefore the higher level phenomena add something extra to the ontology of the world. They each bring a novel structure as well as new causal powers that work downward – going beyond the causal powers of the basal conditions on which they depend.[4] Kim targets the notions of irreducible dependence as well as downward causation in his criticisms.

Though Kim has strong objections to the notion of new causal powers downwardly influencing their basal conditions, he sets these concerns aside in the article I am considering in this chapter. He instead focuses criticism on the notion of higher level phenomena being dependent (i.e. supervenient) on their lower level bases while simultaneously being metaphysically unpredictable given knowledge of those lower-level/basal conditions.[5] When Kim uses the 'supervenience' label to describe the nature of this high-level to low-level dependence, he conceptualizes it as upward determination. He asserts this conception is legitimate with respect to strong emergence based on C. D. Broad's description of the relationship between emergent living bodies and their underlying nonliving chemical compounds.[6] With this particular notion of 'supervenience-as-upward-determination' in hand, Kim characterizes strong emergence

---

[4] Kim, 87.

[5] Kim explores objections to downward causation throughout *Essays in the Metaphysics of Mind* (see that title's index on p.312); however, his first objections are provided early on in the chapter "Making Sense of Emergence" on pages 25-40. I will explore downward causation in detail in Chapter 2 of this paper.

[6] Kim, 88.

as involving: *the claim that some higher level phenomena are not deducible from their basal conditions even though they are supervenient on and determined by them.*[7]

In contrast, higher level phenomena that are *not* emergent from their originating basal conditions are often claimed to be ‘mechanistically explained’, whereby the higher level is simply ‘deducible’ from the lower. Kim sees an entwinement of the concepts ‘explain’, ‘reduce’ and ‘deduce’. He treats reduction as a kind of explanation, and he sees a two way relationship between reducibility and deducibility. In his words, “It is widely assumed that there is an important connection between logical deduction on the one hand and explanation and reduction on the other. Thus, we might add: F is mechanistically explainable, or reducible, in case F is deducible from facts about the constituents of a whole.”[8]

Given this view, we have three contrasting ways to evaluate the relationship between higher level phenomena and lower level phenomena, whereby only two of those ways are characterized as ‘emergence’ by Kim:

1) Mechanistic – the higher supervenes on the lower and is deducible from the lower;

2) Weak Emergence – the higher supervenes on the lower and is not deducible from the lower due to epistemic limitations;

3) Strong Emergence – the higher supervenes on the lower and is not deducible from the lower due to metaphysical impossibility.

[7] Kim, 90.

[8] Kim, 91.

Note that weak emergence – as Kim describes it – is metaphysically equivalent to mechanistic dependence in that deduction of the higher from the lower is *possible* – we just lack the knowledge and/or ability to do it. This leads him to conclude weak emergence claims amount to unremarkable platitudes. In contrast, his subsequent claim of incoherence regarding strong emergence hinges on whether the upward determination of supervenience is a logical entailment or, instead, a nomological (i.e. law like) determination. He asserts it is logical entailment and this leads him to conclude the concept of strong emergence is incoherent. Following is one way to summarize the intent of his argument:

> *P1. Strong emergence includes the claim that some higher level phenomena are supervenient on and yet not deducible from their basal conditions.*
>
> *P2. The type of supervenience involved is logical supervenience.*
>
> *P3. Logical supervenience is entailed upward determination.*
>
> *P4. Entailed upward determination is deducibility.*

> *C1.Therefore strong emergence is the claim that some higher level phenomena are deducible from and yet not deducible from their basal conditions.*

Since P3 and P4 are identity statements equating logical supervenience to deducibility, we can alternately state C1 in a form that more closely matches the wording of Kim's incoherence assertion:

> *C2. Strong emergence is the claim that some higher level phenomena are logically supervenient on and yet not logically supervenient on their basal conditions.*[9]

1.2. Against Kim's Claim of Incoherence:

Kim acknowledges that emergence advocates might want to attack his premise 2 – i.e. strong emergence must involve *logical* supervenience. The emergentists might prosecute their attack by first noting that Broad posited trans-ordinal laws linking basal levels to emergent levels. Therefore, the dependence of higher levels on lower is nomological and not a matter of logical entailment. Kim counters this possible line of attack by asserting the deduction base and the supervenience base *must be held identical*. In other words, if we are claiming we cannot deduce the higher from the lower and yet the lower is supervenient on the higher, then *all* of the conditions in the lower which determine supervenient dependence must be included in the group *from which* we claim we are unable to deduce. In other words, Kim is first asserting that Broad cannot reasonably claim the higher depends on the lower without involvement of his putative trans-ordinal laws. That is, the emergence of the higher depends on the lower level conditions *inclusive* of those trans-ordinals. The trans-ordinals are a necessary condition for the higher level to emerge. When an emergentist then further claims that the higher

[9] Kim, 101.

cannot be deduced from the lower, they cannot take the trans-ordinal laws away and exclaim, "See! I cannot deduce the emergence of the higher." If they do so, Kim can reply, "Then neither can you claim supervenience. So where is your emergence?" His point is that the posited trans-ordinal laws are (under Broad's view) meant to necessarily entail the higher level comes about from the lower. As such, the claimed supervenience is logical entailment. Therefore P2 holds, and strong emergence is – according to Kim – an incoherent concept. Is he correct?

Perhaps, like me, you smell something fishy. According to the terms under which he has framed the argument, Kim is correct. However, it seems clear that he has framed the argument so as to talk past Broad and the other strong emergentists – as if there were nothing in emergent phenomena for either side to be concerned about. In fact, he even seems to revel in talking past Broad when he writes, "Broad will cheerfully admit that if trans-ordinal laws are admitted as part of the deduction base … emergent properties of phenomena are easily deduced from the basal conditions."[10] Well, perhaps Broad would make that admission; however, I doubt he would do so cheerfully. He would more likely be annoyed that Kim is pretending not to notice the important thing that emergentists want to talk about, namely: the fact that (at least with respect to the relationship between physical brain/body states and phenomenal consciousness) we are forced to posit mysterious laws or facts that are *not* reliably reducible to the natural laws of physics. Nor is the presence or nature of these laws logically entailed. We do have knowledge of rough correlations between brain/body states and associated phenomenal reports; however, this knowledge is arrived at *inductively* not deductively. Kim's incoherence

---

[10] Kim, 101.

claim functions as a red herring to call our attention away from the mysterious irreducible nature of, for example, the correlations we have noticed between brain/body states and phenomenal/experiential states.

Broad and his fellow emergentists can avoid Kim's claim of incoherence by pointing to the imprecise correlations between brain/body states and phenomenal states and asking, "Where is the logical entailment? We see only empirically-discovered, rough correlations with many exceptions." Those imprecise correlations are part of the motivation beneath the special sciences (e.g. psychology, macro finance, etc.). Along these lines, Jerry Fodor has called our attention to the fact that special science laws are only generalizations.[11] For example, the predicates associated with different psychological states appear to be analyzable via rough correlations between different mental states. These correlations lead psychologists to posit special science laws of psychology. Beneath each of the mental state predicates is a *wide disjunction* of physical states that could lead to or correspond with a given psychological state. Fodor refers to the connections between each of these disjoined physical states and the mental state that they correspond to as 'bridge laws', and these bridge laws have kinship with Broad's trans-ordinals.

Suppose we posit a psychological law such as, for example: experiencing insults leads to experiencing anger. Beneath each of the predicates (i.e. 'is insulted', 'is angry') there exists a wide disjunction of many different physical states that could underlie being insulted or being angry. Each of those different physical states connects to its mental

---

[11] Jerry Fodor, "Special Sciences (Or: The Disunity of Science as a Working Hypothesis)" *Emergence, Contemporary Readings in Philosophy and Science,* editors Mark A. Bedau and Paul Humphreys (Cambridge: The MIT Press, 2008) 395-409.

predicate via its own distinct bridge law, and each of the physical states for 'is-insulted' has a causal connection to a resultant physical state that, in turn, bridges to 'is-angry'.[12] My point is these widely disjunctive 'laws' are imprecise inductions and not entailed deductions.

In addition to pointing out the widely-disjunctive, multiple-realizability of the higher level states that are studied by our special sciences, Fodor importantly calls attention to the fact that the bridge laws reaching up from the physical to the mental are *not* exceptionless. Think of, for example, political correctness and the many different ways that we must consider how an utterance may or may not be insulting to some people while not to others. Compound that with the realization that, even for a given individual, hearing the same utterance can cause insult on some days and under some circumstances but not on others. Fodor is noticing how the laws of special sciences like psychology have exceptions while the laws of physics do not. He concludes we must accept that the purported bridge laws are not exceptionless and therefore, reductivism with respect to the special sciences "loses its ontological bite".[13] We must therefore accept that these 'laws' cannot support entailment claims – i.e. claims like physical state X entails mental state Y. Therefore, we may look upon Fodor's argument as good reason to deny Kim's claim that the supervenience of emergence is a claim of logical entailment.

Another way to deny P2 – the assertion that emergence involves *logical* supervenience – is to consider how we might arrive at a solution to fill in the mind-body

---

[12] Fodor, see p. 404 for discussion and a useful diagram of these relationships between disjunctions and bridge laws.

[13] Fodor, 405-6.

explanatory gap. For example, Terry Horgan analyzes emergence by first considering the shortcomings of materialism. He defines materialism as holding that:

1. The actual world is a minimal physical duplicate of itself, and
2. Minimal physical duplicate worlds must be identical in all intrinsic respects, and
3. There are no brute/inexplicable inter-level relations linking physical and nonphysical properties or particulars.[14]

Claims 1 and 2 ensure that mental properties and particulars are completely entailed by the physical. Claim 3 ensures that the mental is reducible to the physical. The combination of all three requirements leads to a seemingly insoluble explanatory gap between physical and mental. That is, our most detailed descriptions of brain/body states do not end with '…and therefore E is entailed' whereby E is an experiential/phenomenal property like pain, joy, anger, etc. Yet we know, first hand, that pain, joy, anger and other mental phenomena exist and that we can empirically measure approximate physical correlations beneath them. As physical science progresses, those approximations become more and more precise. However, there seems to be no level of precision that will turn the correlation into a logical entailment.

Horgan suggests we may want to use *abductive* reasoning – i.e. reasoning to the best explanation – to determine how we might arrive at a solution. We might reason that the best way to explain our epistemological explanatory gap is to allow that it "…reflects an underlying *metaphysical* explanatory gap, constituted by the fact that there are

[14] Terry Horgan, "Materialism, Minimal Emergentism, and the Hard Problem of Consciousness" *The Waning of Materialism* editors G. Bealer and R. Koons (Oxford: Oxford University Press, 2010) 309-329.

metaphysically brute necessitation relations between certain physical or functional properties on one hand, and associated phenomenal properties on the other hand—either brute inter-level *laws*, or brute inter-level relations of *metaphysical necessitation*."[15]

Under this view, the supervenience of emergent mental phenomena on physical properties and particulars is a matter of brute metaphysical necessity. However, this conclusion was arrived at abductively, and it is defeasible for the very reason that it was *not* arrived at via *logical* deduction. Therefore we have another coherent view of strong emergence that does not involve logical supervenience.

Since there are at least some conceptions of strong emergence that do not involve logical supervenience, Kim's P2 is false and the concept of strong emergence is not necessarily incoherent. However, concluding that strong emergence is not an incoherent concept and concluding that it actually exists are two different matters. I am uncertain as to what would constitute an existence proof. On the other hand, as long as there are unexplained higher level properties with rough correlations (instead of genuine reductions) to basal properties, it is unclear how we would prove strong emergence does *not* exist.

Any existence proof for strong emergence would have to overcome objections to the additional controversial property of downward causation. Since Kim has put that issue aside for the essay under consideration, I will remark on it only briefly in this chapter and consider it more carefully in the next. Downward causation deserves more careful

[15] Horgan, 320.

consideration because Kim has elsewhere offered compelling arguments against it on the basis of over-determination.[16]

Here is an example of the over-determination worry: If all physical events are caused by other physical events, then the brain/body states that cause conscious phenomenal states must also be the cause of any resultant physical responses. Touching a hot object causes brain/body states that, in turn, cause avoidance action. If so, the phenomenal experience of pain is excluded from the causal chain of the avoidance behavior since the avoidance behavior's cause is already accounted for. Therefore, claims of mental causation (e.g. the experience of pain caused the avoidance) over-determine our actions. The over-determination is due to the apparent presence of two distinct causal chains leading to the avoidance act. If the physical causal chain of brain/body states can fully explain the act, the phenomenal experience of pain is not needed to explain it. On the other hand, it seems obvious to those of us who have touched hot objects that our subsequent avoidance behavior was due to our experiencing pain.

Karen Bennett offers a detailed discussion of the claims and counter claims associated with over-determination.[17] Her preferred strategy to avoid over-determination argues that physical needs mental in order to cause action yet mental does not need physical.[18] In other words, only one of the dual causal chains is necessary to determine

---

[16] See Section V of Kim's "Making Sense of Emergence" in *Essays in the Metaphysics of Mind* (Oxford: Oxford University Press, 2010) 8-40.

[17] Karen Bennett, "Why the Exclusion Problem Seems Intractable and How, Just Maybe, to Tract It", *Nous*, 37:3 (2003) 471-497.

[18] Bennett, 490.

the act, and that one involves the mental. There is a sense in which this seems true. We can imagine someone having a local anesthetic injected into her hand, whereby her nerve endings are temporarily deadened. If she were to come into contact with a hot stove burner, she might not remove her hand immediately since she would not have the phenomenal experience of pain. Therefore, it seems physical injury needs the phenomenal experience of pain to cause the avoidance behavior. Yet the mental/phenomenal might not need the physical since there are cases when we momentarily think we feel a pain and we move only to realize that there was nothing injuring us. Additionally, there are many cases whereby we make decisions to move without any physical stimulus being the cause. For example, I can mentally decide to blink three times in rapid succession without any exterior physical cause other than my interior mental volition – or so it seems.

Despite its seeming plausibility, I am not convinced that the anesthetized-hand succeeds in overcoming the over-determination complaint. Nor am I convinced that Bennett's more general invocation of 'physical requires mental but not vice versa' is successful. My doubt persists because the particular physical instantiations that commonly raise the over-determination worry are neurological. That is, mental states show strong empirical correlation to physical neurological states. It is difficult for me to imagine how the mental cannot rely on those physical states or, at least, some equivalent physical surrogates. That is, any act involving mental choice requires these physical neurological states, and therefore there is no sense in which the mental does not need the physical causal chain riding beneath a volitional act. Therefore the question of strong emergence with downward causation is, in my mind, still open to the worry regarding causal over-determination.

Nonetheless, if we are willing to put the over-determination worry to one side and focus only on Kim's claim of incoherence, we can – for the reasons I have provided earlier – reasonably assert that strong emergence is *not* an incoherent concept. In the remainder of this chapter I will argue that, even if it were to turn out that no solution is available for the over-determination problem (and strong emergence therefore remains in doubt) there is still something worthy of interest in the conceptual space between Kim's notions of strong and weak emergence.

1.3. What About Weak Emergence?

Where Jaegwon Kim sees only two types of emergence, Mark Bedau sees three.[19] He labels them 'Nominal', 'Weak', and 'Strong'. Yet his notion of weak emergence is very specific whereas Kim's is more general. I will describe Bedau's weak emergence after I have first described his notions of the two emergence types that bracket it.

Bedau's nominal emergence is simply one of "…a macro property that is the kind of property that cannot be a micro property." In other words, "… [nominally] emergent phenomena are autonomous from underlying phenomena in the straightforward sense that emergent properties do not apply to the underlying entities." [Bracketed content added for clarity.][20] As Bedau points out, the sheer breadth of nominal emergence is its

---

[19] Mark A. Bedau, "Downward Causation and Autonomy in Weak Emergence", *Emergence, Contemporary Readings in Philosophy and Science,* editors Mark A. Bedau and Paul Humphreys (Cambridge: The MIT Press, 2008) 155-188.

[20] Bedau, 158

weakness. Mere resultants can be predicted and explained by their underlying constituents, and yet they often have properties that do not apply to their underlying causal entities. Therefore ordinary causal resultants, whether understood mechanistically or partially shrouded by epistemic limitations, qualify as nominal emergents.

Just as Kim complains that his own notion of weak emergence is uninteresting and platitudinous, we may view Bedau's nominal emergence in the same light. I see only a small distinction between Kim's notion of weak emergence and the phenomenon that Bedau is labeling 'nominal emergence'. For the remainder of this work I will treat them as the same concept for the sake of convenience. Even so, it is perhaps unfair to claim that mere resultants are philosophically uninteresting. Their status in causality is still worthy of wonder with respect to locating the source of their seeming necessity. As Hume wrote,

> "When we look about us towards external objects, and consider the operation of causes, we are never able, in a single instance, to discover any power or necessary connexion; any quality which binds the effect to the cause, and renders the one an infallible consequence of the other. We only find that the one does actually, in fact, follow the other."[21]

Following Hume, the account of causality still seems somewhat incomplete. In that respect, both Bedau's nominal emergents and Kim's weak emergents are not entirely uninteresting. However, our ability to predict them with great regularity is, perhaps, why

---

[21] David Hume, *An Enquiry Concerning Human Understanding* 1748, *Central Readings in the History of Modern Philosophy, Descartes to Kant,* Second Edition, edited by Robert Cummins and David Owens (Belmont, CA: Wadsworth Publishing Company, 1999) 351.

we do not see them as mysterious. Nonetheless, I put them aside for the remainder of this chapter in my drive to focus on the more controversial aspects of emergence claims.

At the other end of the emergence spectrum, Bedau's strong emergence coincides with that of Kim. In addition to possessing properties not present in their underlying/basal conditions, strong emergents are claimed to be supervenient on and yet irreducible to their bases. Like Kim's strong emergence, the emergent properties in Bedau's notion are thought to have new causal powers that are efficacious at both the emergent level and the underlying basal level.[22] Thus Bedau's notion of strong emergence is, as near as I can tell, the same as Kim's.

Bedau acknowledges the usual worries associated with strong emergence in regard to downward causation and over-determination as well as violation of causal fundamentalism.[23] He also seems to possess Kim's lack of interest in the ordinariness of nominal emergence. Yet, unlike Kim, Bedau sees something interesting in between the ordinariness of nominal emergence and the controversial conception of strong emergence. He puts forward an in-between conception that he (unfortunately) labels 'weak emergence'. In order to avoid confusing Bedau's weak emergence with Kim's, I will refer to Bedau's as 'B-weak emergence'.

According to Bedau, B-weak emergents are *ontologically* dependent on and reducible to their micro bases. They are also *causally* dependent on and reducible to those same bases. Even so, these B-weak emergents have a kind of autonomy that is

---

[22] Bedau, 158-159.

[23] Bedau, 159.

explanatorily irreducible due to a particular type of complexity.[24] Taken together, these attributes may seem contradictory in that they posit ontological reducibility along with causal reducibility while also positing explanatory irreducibility. We often think of ontological and causal reduction as amounting to explanation. However, there is a sense in which these reductions are not a *full* explanation. This becomes apparent when we, for example, consider mental phenomena. It seems clear that even if we could formulate a causal reduction of consciousness in physical and bio-chemical terms, we would not have a full explanation since physical reductions do not explain the phenomenal aspects of what it is like to experience the world via consciousness. This is made clear by Frank Jackson's 'Mary' thought experiment. In that experiment, Mary has never seen color and yet is aware of all the microphysical facts and explanations as to how vision works, how objects emit or reflect colored light, and how color is perceived. If physical reductions fully explain color, Mary ought to know what it is like to experience red even though she has never done so. Yet our intuitions tell us that she would not know what it is like to experience red unless she had actually done so. Having grown up in a black-and-white world, she would likely gasp with excitement the first time she actually saw a red rose.[25]

To be clear, Bedau does not claim B-weak emergence helps explain the mind/body gap. In fact, he explicitly states it would be an "unsatisfying" tool to explain phenomenal consciousness.[26] Nonetheless, the mind/body problem is a good illustration as to how we can have physical causal reducibility with explanatory irreducibility. In the less-

[24] Bedau, 160.

[25] Frank Jackson, "Epiphenomenal Qualia", *The Philosophical Quarterly*, Vol. 32, No. 127 (Apr., 1982), pp. 127-136.

[26] Bedau, 161.

ambitious case of B-weak emergence, Bedau asserts that this phenomenon of causal reducibility with explanatory irreducibility is the result of 'computational irreducibility' – i.e. not predictable by any algorithmic method. He arrives at this conclusion because B-weak emergence pertains to macro properties that are underivable *except* by simulation. More specifically: If 'P' is some property that is nominally emergent from some system 'S', P is B-weak emergent (in addition to being nominally emergent) "…if and only if P is derivable from all of S's micro facts but only by simulation."[27]

This 'derivability-by-simulation-only' is not merely an epistemic limitation. It is also a formal limitation. According to Bedau, derivation of a B-weak emergent by any means other than simulation is not possible – even for a super-calculating archangel or a "Laplacian supercalculator".[28] By 'simulation' Bedau means aggregation and step-by-step iteration of the base micro facts and properties or (presumably) sufficient representations of those micro facts and properties. Additionally, Bedau importantly notes that, under his notion of B-weak emergence, there is no possible simulation of a B-weak emergent that is simpler than the intrinsic natural process by which the system behaves.[29] Thus any computational simulation of a B-weak emergent process will be more complex than simply aggregating the real micro-constituents and watching it unfold. This is why Bedau refers to B-weak emergents as 'computationally irreducible'.

---

[27] Bedau, 162.

[28] Bedau, 163. Bedau's use of the term 'Laplacian supercalculator' is in reference to a super-computing being imagined by Marquis de Laplace as having the capacity to predict the future based on complete knowledge of all present microphysical facts. Pierre Simon the Marquis de Laplace, *A Philosophical Essay on Probabilities* translated by Truscott and Emory (New York: John Wiley & Sons, 1902) 4.

[29] Bedau, 164.

Bedau invokes cellular automata and variations on John Horton Conway's Game of Life to illustrate B-weak emergence.[30] The Game of Life uses metaphorical labels to describe emergence of unexpected macro-level behaviors in certain computer simulations operating under fixed, micro-level rules of display transitions. Imagine a 'universe' as simple as an unlimited two-dimensional plane divided into square cells. Cells that are 'alive' are black. Cells that are 'dead' are white. Such a universe is portrayed along with a single living cell in the following diagram.

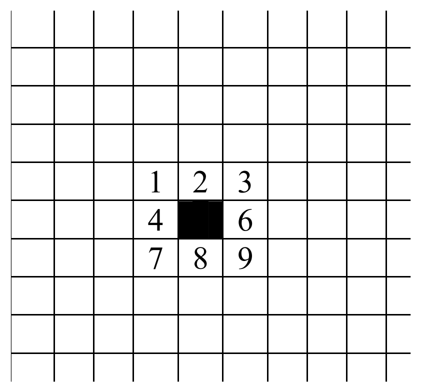

Figure 1. Example Game of Life Universe

From the diagram, cell 5 is alive, and all others are dead. The cell numbers are arbitrary. With each tick of the clock in this universe, the status of each cell is updated according to some defined function. Researchers have applied many different update functions to this sort of cellular automata universe; however, Bedau and others attribute the most famous update function to John Conway.[31] It is called the 'Game of Life' and it delivers some remarkable macro-level behaviors.

[30] For a more in-depth review of Conway's Game of Life see Cornell's "Math Explorer's" website at: http://www.math.cornell.edu/~lipa/mec/lesson6.html (Visited 12/16/12)

[31] Bedau, 166.

Game of Life Update Function:

A living cell remains alive if and only if either two or three of its neighbors were alive at the previous moment (i.e. the previous clock tick), and a dead cell becomes alive if and only if exactly three of its neighbors were alive at the previous moment.[32]

Given this update function cell 5 in Diagram 1 will die with the very next clock tick. Nonetheless, there are many different initial configurations of living cells that result in continued life, and this continued life evolves in remarkable patterns that *cannot be determined or calculated in advance by any method simpler than running the simulation.* By 'running the simulation' I mean letting the clock tick, applying the update function with each tick, and observing what happens. What happens will, of course, depend on the initial conditions as well as the update function. Some initial conditions evolve into a completely dead universe. Others evolve into still life forms that make no further changes. Still others evolve into oscillating forms that repeat the same visual sequence over and over within a localized area. Still others evolve into apparently endless chaotic expansions of seemingly unpredictable patterns moving out into and filling the universe.

One particularly interesting class of patterns are labeled 'gliders' because they seem to glide across the plane on a diagonal, and they do so indefinitely. I've illustrated two gliders in the diagram below. With a little mental effort you can apply the Game of Life update function and a few ticks of the clock to these patterns. As the clock ticks the gliders shape-shift slightly and then return to their original shape only to be rotated and

[32] Bedau, 166.

advanced diagonally. Under slow clock sequences they appear to waddle off into space diagonally. When the clock ticks fast, they appear to glide.

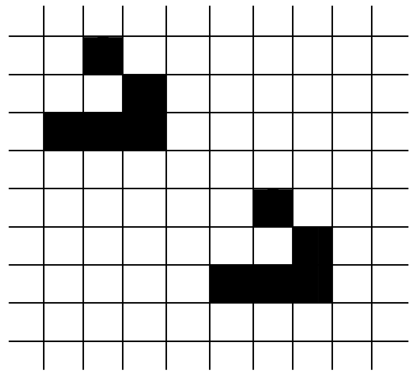

Figure 2. Two Gliders

We can form gliders as part of an initial configuration in a game-of-life universe and watch them glide; however, this fact is not particularly remarkable on its own. The remarkable fact is that there exist a number of different initial configurations of cell structures that *produce* streams of gliders. Experimenters have discovered these 'glider gun' configurations quite by accident through *simulation*. We don't know how many different glider gun configurations there are since there is no computational algorithm to predict them – at least none that is any simpler than configuring a candidate gun (or precursor) and running the simulation. In this way, continuous glider guns qualify as B-weak emergents that are multiply realizable via an indeterminate number of different gun configurations. The glider guns – as a class – are computationally irreducible precisely because there is no computational algorithm that can predict how many glider gun configurations exist, nor even how to go about finding the next configuration.

In addition to glider guns with glider streams, experimenters have discovered a large number of other novel B-weakly emergent structures in the Game of Life. A web search for ‘Conway Game of Life’ will reveal many available programs allowing users to run their own experiments.[33] It is also worth noting that experimenters have discovered B-weak emergent patterns in other cellular automata universes using update function rules different than Game of Life.[34]

Cellular automata, the Game of Life, and glider guns exemplify and distinguish B-weak emergence from mere nominal/mechanistic emergence – i.e. from Kim’s weak emergence. Instances of these B-weak emergent phenomena are multiply realizable and have clear ontological and causal reduction paths back down to the initial, lower-level conditions – i.e. the initial cell configurations and update functions. Yet, unlike mere resultants, there is no corresponding deduction path that leads us up to the discovery of the full range of all glider gun configurations. We can, of course, write computer programs that will try every combination of live-vs.-dead initial cell configuration for a given area of the universe. Such a program could tell us whether a glider gun is discovered. Nonetheless, this method is not discovery by deduction. It is trial-and-error simulation, and it only covers the particular initial cell arrangement area considered – as opposed to the infinite number of possible areas and initial configurations. Thus, in the Game of Life, we have the emergence of higher-level, novel, objective structures (i.e. the class ‘glider guns’) from lower-level phenomena (i.e. cells and update functions), and –

[33] For a particularly good Conway Game of Life simulation that you can run on a desktop computer see: Paul Callahan’s “What is the Game of Life?” http://www.math.com/students/wonders/life/life.html (Accessed 12/16/12) or Edwin Martin’s “bitstorm.org”, John Conway’s Game of Life, http://www.bitstorm.org/gameoflife/standalone/, (Accessed 3/3/13). Also, you can download a 3000+ archive of Game of Life patterns at the LifeWiki, , http://conwaylife.com/wiki/Main_Page, click on “Download pattern collection”, (Accessed 3/3/13).

[34] Bedau, 170-171.

counter to Kim – the instances of these emergent structures are ontologically and causally *reducible to* their basal conditions though the class is not *deducible from* those bases.

1.4. A Challenge to B-weak Emergence and a Reply

At this point, a Kim advocate might object by pointing to any given instance of a glider gun and claiming, "I *can* deduce that this configuration of living cells operating under Conway's update function will produce a stream of gliders." He would then produce a detailed description of the logical entailment for each next-clock-status of every cell based on the function rules and perhaps reminding us along the way of the higher level appearance that these logical entailments will generate – i.e. gliders moving diagonally. We might respond by complaining, "You are merely running the simulation with your mind and voice!" Yet our challenger could fairly reply, "I am performing a logical deduction of your putative novel, irreducible structure and it looks like a mere resultant to me!" So who is correct?

At this limited level of analyzing a single glider gun instantiation, both are correct; however, Bedau is not limiting his observations to individual instantiations of glider guns. He is noticing (correctly I think) an overarching autonomous structure with respect to glider stream origins. *Glider guns are multiply realizable via an indefinite number of configurations.* Because of this indefinite multiple-realizability, it is not enough to reduce a given instance of a stream of gliders to a particular gun under consideration. This reduction – and the opposite-way corresponding deduction that the gun will shoot gliders – is *not a complete explanation* of glider streams and guns. The explanation of the existence of emergent glider streams and the sorts of things that emit them remains an autonomous feature of the Game of Life. In Bedau's words,

> "...weak emergent phenomena that would be realized in an indefinite variety of different micro contingencies can instantiate robust macro regularities that can be described and explained only at the macro level. The point is not just that macro explanation and description is irreducible, but that this irreducibility signals the existence of an objective macro structure."[35]

In the case of Game of Life glider guns and their associated glider streams, the 'objective macro structure' that Bedau is referring to is *not* the thing you see when you examine a particular instance of a glider gun spewing gliders. It is the emergent phenomenon of glider gun-ness in general – i.e. in all of its various multiply realizable configurations. When we take a given stream of gliders marching from a gun and reduce that stream to that gun, we have not reduced glider streams to glider guns in general. If we had no knowledge of the particular gun at work and only had visual access to the stream of gliders, we could not deduce what was actually causing the stream since glider guns are multiply realizable in an indefinite number of configurations. Thus glider gun-ness is a robust form of B-weak emergence whose instances are individually causally and ontologically reducible to their basal conditions, but whose macro-level phenomenal structure is autonomous in that it is not fully explained by reduction of instances – and also not fully deducible from the update functions of the Game of Life as evidenced by the indefinite number of ways for glider guns to occur. As Bedau concludes, "This kind

[35] Bedau, 182-183.

of robust weak emergence reveals something about reality, not just about how we describe or explain it. So the autonomy of this robust weak emergence is ontological, not merely epistemological."[36]

Bedau has illuminated a type of emergent macro structure having an irreducible nature despite its individual instances being both causally and ontologically reducible. Thus, counter to Kim, Bedau has explicated a coherent notion of irreducibility in the face of supervenience. This type of B-weak emergence is not full-blown strong emergence since the causal powers associated with the higher-level emergent phenomena do not compete with lower level causes in ways that raise the issue of over-determination. On the other hand, this type of B-weak emergence is not as simple as the mere resultants of nominal emergence in that – riding atop the causal explanations of individual instances – we have autonomous macro-level phenomena. These macro-level phenomena are not explained across their full and indefinite number of possible realizing structures by examining any given instance or group of instances. Thus Bedau has explicated a coherent kind of emergence occupying the conceptual space between Kim's notions of strong and weak emergence. Furthermore, the autonomous, ontologically objective nature of these emergent macro structures is worthy of our philosophical interest despite the mechanistic reducibility of their individual instantiations.

1.5. Concluding Thoughts on B-weak emergence and One More Concern

Multiple realizability in combination with computational irreducibility help to define and make interesting Bedau's notion of B-weak emergence. Together they signal an

---

[36] Bedau, 183.

objective, overarching, higher-level phenomenon that is more than a mere resultant. Nonetheless, we can question whether Bedau has presented *decisive* reason to support his assertion that the autonomy of B-weak emergence is ontologically objective and not merely apparent due to epistemological limitations.

Consider again the example of glider guns. As described above, there are many different initial configurations of cellular automata that will produce glider streams, and the total number of different initial configurations with this capability is indeterminate. We have no way of knowing how many there are. This means that, even if we were to discover all of them, we would not know it. Now suppose that, under some other hypothetical update function (call it 'Life-Alt') we notice certain initial configurations produce streams of 'hoppers' rather than gliders, and suppose we discover two such hopper gun configurations. Finally, suppose that those two are the only metaphysically possible hopper guns under Life-Alt rules. Since the cellular automata universe extends infinitely on a two-dimensional plane, there are an infinite number of possible initial configuration combinations that we would need to explore in order to know we had found every possible hopper gun. We would never know the two we have are all there is.

Under the thought experiment I have just described, an antagonist toward B-weak emergence might claim that a mere two initial configurations just happening to spew hoppers is not enough to assert an objective, ontological, higher-level, phenomenon of hopper gun-ness. The two hopper guns could be thought of as mere happenstance instead of signaling a metaphysically objective ontology. Yet, if we accepted Bedau's notion of B-weak emergence, our limited epistemic position in the hopper gun case would force us to conclude we are observing an objective ontological structure since (from that perspective) the total number of hopper gun configurations is indefinite and we've found multiple realizability – albeit only two. Our epistemic limitation with respect to there being only two versus an indefinite number of hopper guns is a matter of metaphysical necessity due to the infinite span of the cellular automata universe and the fact that we possess no algorithm to predict additional hopper guns. However, our conclusion that we are observing an objective, higher-level, ontology is driven, at least in part, by that epistemic limitation, as opposed to there actually *being* a higher-level phenomenon – or so the argument goes.

With this line of thought, I am suggesting that there may or may not exist an objective higher level phenomenon depending on whether we are justified in claiming the presence of only two configurations is pattern-enough to signal something real. In this particular example, it is our epistemic limitation regarding the apparent indefinite number of configurations (combined with our lack of predictive algorithm) that drives us to the B-weak emergence conclusion – and perhaps not the actual existence of a higher-level phenomenal structure.

We might attempt to resist this worry by countering that two configurations *are* enough to establish the existence of an objective, higher-level phenomenon. We only need multiple-realizability in combination with an indefinite number of ways to realize. Two instances qualify as multiple, and a finite being could never discover the fact that there are only two hopper guns in Life-Alt. These facts qualify our two-hopper example as computationally irreducible – or so this argument goes.

Given the above exercise, it seems that the ontological status of B-weak emergent structures has a fuzzy boundary based on our intuitions about frequency of instances. How many different instantiations are enough to amount to a recognizable pattern justifying an ontological claim? Two hardly seems enough. Moving up from two we are faced with a sorites paradox along the lines of trying to decide when one more grain of sand turns a collection of grains into a heap.[37] This is a disappointing feature of B-weak emergence in that we don't want to think of 'exists' as a vague predicate based on an imprecise number of observed realizers.

On the other hand, despite the fuzzy threshold for qualifying as a B-weak emergent, macro structures with *many* realizers do seem to signal a genuine identity that is explanatorily autonomous with respect to the full and indefinite number of its realizers. *A sorites paradox does not challenge the ontological status of heaps – only the qualifying boundary of instances.* Furthermore, in the examples of emergence considered, the individual instances are reducible both causally and ontologically. From this I conclude that we should credit Bedau with identifying a slice of emergent phenomena that overcomes Kim's incoherence objection regarding supervenience with irreducibility while still offering something more than uninteresting mechanistic causality.

In the particular Kim work considered here, the thrust of his complaint amounted to an assertion that it is incoherent to claim a higher level phenomenon is causally dependent on the behavior and properties of its underlying constituents and yet not reducible to those same behaviors and properties. Kim views an emergent's dependence

---

[37] "Sorites" from the Greek "soros" for heap. Simon Blackburn, *Oxford Dictionary of Philosophy, second edition,* (Oxford: Oxford University Press, 2005) 345.

as a logical entailment from the lower level behaviors and properties to the higher level behaviors and properties. Therefore, it is incoherent to claim the reverse direction of reducibility from the higher to the lower does not follow.

Bedau's B-weak emergence avoids this incoherence claim because the individual *instances* of, for example, glider streams and their guns are both entailed by and completely reducible to the cellular automata and Game of Life rules they depend on. Nonetheless, the phenomenon of glider guns *in general* (i.e. glider gun-ness) possesses genuine ontological status because it is multiply realizable via an indeterminate number of many differing structures. This indeterminate realizability is a kind of irreducibility – Bedau's 'computational irreducibility'. Bedau is calling our attention away from the reducibility of individual instances to the overarching phenomenon of gun-ness in general, and he is noticing that, beyond the individual instances, some phenomenal structure exists. That structure is not deductively predicted from, nor reducible to Game of Life rules, even though its individual instances reduce. The indeterminate extent of its multiply realizable structures signals its ontological autonomy, while the reducibility of its individual instances overcomes Kim's complaint of incoherence.

CHAPTER 2

STRONG EMERGENCE AS EMBODIED MIND AND THE WORRY OF DOWNWARD CAUSATION VERSUS EPIPHENOMENALISM

2.1. Kim's Downward Causation Worry

Advocates of strong emergence assert that emergent phenomena have causal powers not possessed by their underlying basal constituents, and emergent phenomena are able to causally influence their basal constituents. Under some views, this downward causation is claimed to be *synchronic* such that the newly emergent causal powers affect the bases as those bases *simultaneously* cause the emergent phenomenon. In other words, the emergent phenomenon supervenes on the very effect it causes in a sort of reflexive, instantaneous, causal circularity. In his book chapter, "Making Sense of Emergence", Jaegwon Kim examines synchronic reflexive downward causation and concludes it is ultimately incoherent because it violates the "causal-power actuality principle" which, according to Kim, is tacitly accepted by most of us.[38] Kim describes that principle as follows.

> "For an object, *x*, to exercise, at time *t*, the causal/determinative powers it has in virtue of having property P, *x* must *already* possess P at *t*. When *x* is caused to acquire P at *t*, it does not already possess P at *t* and is not capable of exercising the causal/determinative powers inherent in P."[39]

In less precise language: cause happens before effect. Even though the interval between the two may be vanishingly small, one must come before the other. Therefore, any talk of synchronic causation might be viewed as incoherent. Some thinkers advocating for synchronic causation might coherently mean that the instantaneous status of a system (ignoring quantum jitters) is the result of prior (and perhaps ongoing) causes feeding back into the system. I believe Kim would (and should) insist that any such state of affairs is not synchronic but, instead, *diachronic* – i.e. occurring across a time interval however small that interval may be.

---

[38] Kim, 35-37.

[39] Kim, 35

Kim also takes issue with diachronic downward causation, and he uses the concept of mental causation to exemplify the problems associated with it. Think of a physical neural state P as the basal condition for an emergent mental state M. Now, consider that we have all had the phenomenal experience of one mental state causing another – e.g. my earlier example of the mental experience of an insult (call it M) causing the mental experience of anger (call it M*). There is strong empirical evidence that mental states have underlying physical neural correlates. In the case of M*, call the underlying neural base P*. Here is a schematic representation.

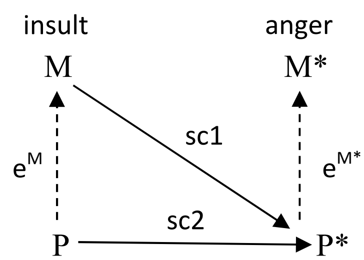

Figure 3. Insult Causes Anger[40]

In Figure 3, e^M represents the emergence of the mental state of experiencing insult M from its basal neurological state P. Similarly, e^{M*} represents the emergence of experiencing anger M* from its basal neurological state P*. Our experience tells us that the insult caused us to be angry. If true, then M provides sufficient cause sc1 for P*. Yet physical and chemical theory supported by empirical evidence indicate that P provides sufficient cause sc2 for P*. From this scenario, Kim concludes that downward causation

[40] This graphical representation is a modified version of a more generalized graphical representation of the causal overdetermination problem. The more generalized version was presented to me in white board illustrations by Bernard W. Kobes and Jeffrey J. Watson during the Spring 2012 seminar "Emergence and Cosmic Hermeneutics" at Arizona State University.

involves "causal overdetermination."[41] P is able to do all the work needed to cause P* and the follow-on emergence of M*. Therefore, M is not needed as the cause. The phenomenal experience of insult is not necessary to cause the anger. The underlying neural correlates are all that is needed. Therefore, the phenomenal experience of insult is merely along for the ride. The neural states underlying the insult caused the neural states underlying the anger, and the anger too is merely along for the ride.

One may be tempted to shrug off this overdetermination as an epistemic puzzle having to do with the necessity and sufficiency of multiple causes; however, champions of strong emergence are not afforded that opportunity. The concept of strong emergence is constituted, in part, by the foundational notion that higher-level, emergent phenomena possess new causal powers. In particular, "...emergents are supposed to make distinctive and novel causal contributions."[42] According to the overdetermination objection, emergents make no distinctive causal contribution over and above that of the underlying basal physical contributions.

2.2. Distinguishing Cause and the Experience of Cause

Like me, you might find this epiphenomenal[43] result disconcerting. We have all been unfairly insulted by a fellow human, and have often felt angry as a result. At those times nothing could be clearer to us than the fact that the insult caused the anger – seemingly we had a genuine experience of the causal effect of the insult leading us to the

[41] Kim, 83.

[42] Kim, 83.

[43] Blackburn, 117. "**epiphenomenalism** The view that some feature or situation arises in virtue of others, but itself has no causal powers."

anger. There must be something obviously wrong with Kim's view of the situation. Our experience ought to count as supporting evidence for mental causation. Yet, perhaps our experience misleads. Here is one way to view what might be happening when we experience one mental state as the cause of another.

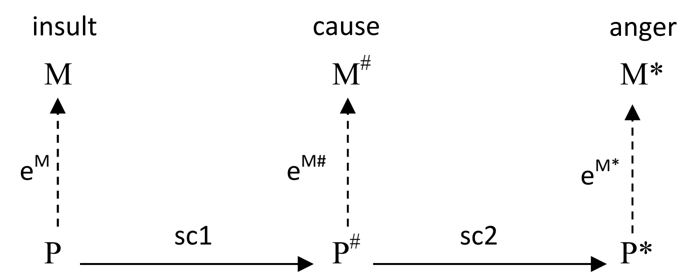

Figure 4. The Experience of Cause

According to the view I present here, the emergent experience of insult does not definitively cause the emergent experience of anger merely because we experience it as the cause – any more than our experience of the earth as being flat is definitive evidence of it being flat. If we experience insult-causing-anger, and we also accept the overwhelming empirical evidence regarding the neural correlates of consciousness, then why not conclude that the experience of insult-causing-anger consists in the phenomenal experience of the neural activity *in between* the fully formed neural state of insult and the fully formed neural state of anger? In other words, neural causation can be experienced just as the neural states beneath insult and anger can be experienced. *All three* experiences (M, $M^{\#}$, and M*) could ride atop their neural bases (P, $P^{\#}$, and P*) without possessing any causal efficacy themselves. If so, the experience of causation is just that – i.e. experience. The causation is occurring in the neural correlates in between the insult and the anger – at least according to this particular view.

To be clear, Kim has not argued that mental causation is experiential and epiphenomenal as part of his polemic against downward causation. My point in presenting the argument is as follows. Our intuition that mental causation exists is driven – at least in part – by our having experienced an apparent causal relation between our own mental states. We cannot safely use that experience to conclude mental causation – viewed as post-emergent downward causation form mental to physical – has genuine ontological status.

At this point in my argument, I am not endorsing the epiphenomenal view presented above in Figure 4. On the other hand, I think it a useful tool to help organize our thoughts on the topic of mental causation. I will reconsider endorsing it after we have explored Evan Thompson's counter to Kim's downward causation worry. In the meantime, think of it as an exploratory tool. By representing our experience of mental causation as located temporally in between consecutive mental states and assuming that all experience has physical correlates, we can examine the causal relations between those physical correlates without worry that the experience of mental causation will disappear. The question remains whether that experience is merely epiphenomenal or genuinely causally efficacious. If it is genuine, how can we conceive of it as being so while avoiding Kim's overdetermination worry? Again, Figure 4 is a useful conceptual tool in this exploration. Notice how Figure 4 and Figure 3 both express the debated positions in a mind/body dualist framework. In both Figure 3 and Figure 4, mind is shown as being apart from and above physical matter. Now suppose that the categorical differences between physical and mental are less pronounced than we intuit. We might represent this as the dotted arrows of e^{M}, $e^{M\#}$, and e^{M*} shrinking in length. When the arrows become length = zero, mind and neural states become identities, and mental causation *just is* physical causation. The problem with this embodied-mind line of thought is that, when arrived at from the direction I have outlined, the notion of emergence has disappeared and we are left with an isomorphic relationship between the physical and mental, namely: $M=P$, $M^{\#}=P^{\#}$, and $M^*=P^*$. This is not a very satisfying result in that a complete description of the physical entities constituting our bodies – including their spatial, temporal, and energetic relationships – fails to yield an explanation of the phenomenal experience we associate with mind.

Is there a way to describe mind as embodied without that description distilling to a reductive isomorphism between mental and physical? Does this description include emergence? Evan Thompson believes the answer to these two questions is yes. In the following subsections I will explore his view.

2.3. Overview of Thompson's *Mind in Life*:

In his recent book, *Mind in Life,* philosopher Evan Thompson sets out to harmonize the experimental sciences of life and mind with the phenomenology of experience and subjectivity.[44] In doing so he hopes to make a contribution toward resolution of the mind-body explanatory gap. Summarized broadly, Thompson views mind as a self-organizing, emergent phenomenon constituted not just in the brain, but reflexively between body and environment – each shaping the structure of the other. As such, mind is not reducible to the neural activity of brains or even entire nervous systems. I will explain these concepts in greater detail below; however, at this stage it is interesting to

[44] Evan Thompson, *Mind in Life - Biology, Phenomenology, and the Science of Mind* (Cambridge: The Belknap Press of Harvard University Press, 2010 paperback, 2007 original)

note that, in sympathy with the views of biologist-philosopher F. J. Varela, Thompson argues the incipient stage of cognition is located in the self-organizing structural processes of individual living cells.

To conceptually locate his view, Thompson contrasts three major approaches to the study of mind since the waning of 1950's behaviorism, namely:

1. Cognitivism – using the digital computer as a metaphor for mind,
2. Connectionism – describing mind as a neural network, and
3. Embodied Dynamicism – mind as an embodied, self-organizing dynamic system.[45]

He describes cognitivism as having replaced the behaviorist practice of treating mind like a black box that receives inputs from the environment and outputs behaviors. Instead, cognitivism uses a computational metaphor. Mind involves computational hardware in the form of a brain and nervous system that runs software amounting to representation processing. The functionalist version of cognitivism holds that the physical substrate doing the processing is irrelevant. It need not be biological. In Thompson's view, cognitivism suffers the shortcoming of failing to explain how brains have experiences and failing to explain the relationship between computation and phenomenal experience.[46]

Connectionism is similar to cognitivism but uses a somewhat modified metaphor of mind as being a neural network. Both cognitivism and connectionism are computational

[45] Thompson, 4.

[46] Thompson, 4-7.

models. Whereas cognitivism posits symbol processing in the brain to achieve reasoning and language, connectionism posits sub-symbolic processing in the form of pattern recognition. In Thompson's view, connectionism also suffers from an inability to account for subjective experience.[47]

He instead favors and promotes the embodied dynamicism view of mind, whereby minds are emergent, self-organizing, dynamic systems such that "... cognitive processes emerge from the nonlinear and circular causality of continuous sensorimotor interactions involving the brain, body, and environment."[48] By 'nonlinear' he means the output of an interaction is not directly proportional to the sum of the inputs – e.g. very small changes in inputs cause disproportionately large changes in outputs.[49] By 'circular causality' he means the global to local interactions of a dynamic system have bidirectional efficacy. In other words, component behaviors influence and generate a global order which influences and constrains component behaviors such that at least some component of the global order is held invariant (i.e. fixed and stable) by the circular process.[50] I think of this as stabilizing feedback.

The above mentioned reflexive involvement of body and environment in cognition is a key distinguishing feature of embodied dynamicism. Mind is not just brains and nervous systems. The reflexive role of body and environment in the formation of

---

[47] Thompson, 9-10.

[48] Thompson, 10-11.

[49] Thompson, 39.

[50] Thompson, 62.

cognition is said to be 'enacted' by way of circular causation. I will offer more detail on this idea shortly.

Thompson gives great credence to the 'enactive approach' of explaining embodied mind, and he credits the concept to his 1991 collaborative work with F. J. Varela and E. Rosch – (i.e. *The Embodied Mind:…*).[51] He likewise gives great credence to the concept of biochemical cellular 'autopoiesis' – a concept he credits to the 1980 work of F. J. Varela and H. R. Maturana.[52] In fact, Thompson's work in *Mind in Life* is largely derived from the enactive approach and autopoiesis. It portrays cognition as a strongly emergent phenomenon.

The enactive approach to explaining embodied mind comprises a collection of five interrelated ideas, namely:

1. Living beings (human or otherwise) are autonomous agents that generate and maintain themselves while enacting their own cognitive domains.
2. Nervous systems are autonomous dynamic systems that *generate meaning* as opposed to processing information computationally.
3. Cognition amounts to exercising skillful know-how in embodied action emerging from recurrent sensorimotor patterns of perception and action.
4. A cognitive being's world is a relational domain brought forth (i.e. enacted) by the being's autonomy and mode of interaction with its environment – as opposed to environment being a pre-specified external realm represented internally.

[51] Thompson, 13.

[52] Thompson, 44.

5. Experience is central to explaining mind – not epiphenomenal.[53]

Thus mind is seen as involving a living being in the form of a dynamic, self-maintaining system that generates its own cognitive world by creating meaning through recurrent sensorimotor interaction. The generation of meaning is a necessary component in the self-maintenance of the system, and the recurrent sensorimotor patterns involved in the generation of that meaning amount to feedback for the dynamic system.

These five ideas revolve around a somewhat different – or at least more broadly construed – view of cognition. *Thompson sees cognition as a kind of know-how associated with embodied action rather than a higher level processing of representations.* In his own words, "Cognition is behavior or conduct in relation to meaning and norms that the system itself enacts or brings forth on the basis of its autonomy."[54] Furthermore, "Something acquires meaning for an organism to the extent that it relates (either positively or negatively) to the norm of the maintenance of the organism's integrity..."[55] In other words, Thompson roots cognition and meaning-giving in the very experience of being embodied and having norms for self-interested action that maintains system stasis.

In this context, cognition at its very lowest level *need not even involve brains, nervous systems or awareness in the sense of wakefulness.* It requires only an organism that is dynamic with feedback to a stasis such that the system behaves according to self-

---

[53] Thompson, 13.

[54] Thompson, 126.

[55] Thompson relating H. Depreester's view of meaning to his own, 70.

preserving norms by selectively reacting to stimuli – i.e. giving meaning. Under this view, Thompson asserts that *all* living things are minimally cognitive.[56]

In support of that assertion, he considers the paradigm example of life in its most basic form – the biologic cell. Cells behave according to a pattern of autopoietic organization. The term 'autopoiesis' derives from Greek 'poiēsis' for 'production'.[57] Autopoiesis is characterized by a system that produces its own boundary (e.g. the cell wall) and self-regenerates via dynamic feedback so as to self-maintain. Self-maintenance means the system is operationally closed; yet it is necessarily thermodynamically open to the outside world since it needs sources of energy for its regenerative self-maintenance.[58] All living organisms possess these characteristics. Again, in Thompson's own words,

> "...the form or pattern of autopoietic organization is that of a peculiar circular interdependency between an interconnected web of self-regenerating processes and the self-production of a boundary, such that the whole system persists in continuous self-production as a spatially distinct individual."[59]

The continuous self-regeneration of the cellular system maintaining its own identity as a distinct entity is seen as a closed operation that regulates to self-enacted norms. If we accept Thompson's broadly-construed definition of cognition as giving meaning to stimuli such that the organism behaves in relation to norms that the system itself enacts

[56] Thompson, 127.

[57] Blackburn, 281.

[58] Thompson, 98.

[59] Thompson, 101.

on the basis of its autonomy, then we are led to Thompson's primary thesis: There exists a deep continuity between life in its most basic form and mind.[60]

Thompson's assertion that *any* living system is cognitive seems difficult to entertain in regard to organisms possessing no brain or nervous systems. It requires we accept his notion of cognition as a certain type of autonomous behavior in relation to self-enacted norms – something well short of processing of disembodied representations.   Just prior to presenting his enactive approach to explaining embodied mind, Thompson lays the ground work to prime our acceptance of his view of cognition by making four important observations:

a) The precise delineation of personal versus sub-personal as well as conscious versus preconscious is unclear.
b) Awareness of one's own psychological and somatic processes varies from subject to subject.
c) Since much (if not most) of what we are as psychological biological beings is unconscious, our subjectivity cannot be understood without relating it to our unconscious processes.
d) Unconscious processes are far from skull-bound neural happenings but extend throughout the body and, via our sensorimotor interactions, loop into and back from the outside socio-cultural environment.[61]

I take Thompson's point in making the above observations to be that, even though modern humans tend to think of mind in Cartesian terms of consciously processing

---

[60] Thompson, 126-7.

[61] Thompson, 12.

representations, this representational processing is (more or less) the tip of the iceberg and available, in part, only to the most advanced cognizers – i.e. humans. It is the prominent feature we focus on while ignoring the larger unconscious or preconscious cognitive processes from which our conscious awareness wells up. The boundary between conscious and pre- or un- conscious is not distinct. Furthermore, cognition transcends that boundary downward into the processes from which consciousness wells up. Thus consciousness in the form of wakeful awareness can be viewed as a higher level of cognition – *but not required for cognition*. How far past the conscious/non-conscious boundary does cognition extend? Thompson views it as extending down to the lowest level of processing constitutive of autonomous behavior to self-enacted norms for self-preservation and production – down to the biologic cell.

2.4. The Role of Emergence in Thompson's View

In the later parts of this chapter I will consider in more detail Thompson's proposal regarding the lower boundary of cognition. At that point we will see how his follow-on primary thesis – that there exists a deep continuity between life in its most basic form and mind – provides an avenue for overcoming *some portion* of the worry over epiphenomenalism with respect to mental causation. Notwithstanding that longer-term objective, my interest in Thompson's work during the early stages of this chapter will focus on the role that emergence plays in his enactive approach to explaining embodied mind and, more importantly, on how he defends emergence against Jaegwon Kim's downward causation criticism. I have presented Thompson's view on cognition because his enactive approach employs emergence in a way that, he believes, overcomes Kim's downward causation worry – or at least makes it seem irrelevant. Therefore, understanding the enactive approach and the underlying supporting role of cellular autopoiesis is important to understanding Thompson's counter to Kim.

Let's consider the role of emergence in the enactive approach. A second look at the five ideas constituting that approach reveals the explicit employment of emergence in idea number three: *Cognition amounts to the exercise of skillful know-how in embodied action emerging from recurrent sensorimotor patterns of perception and action.* In other words, cognition is an emergent phenomenon requiring recurrent interaction with the world. Now recall idea number two: *Nervous systems are autonomous dynamic systems that generate meaning as opposed to processing information computationally*. From this we gather that the recurrent interaction with the world mentioned in idea three is what leads to the generation of meaning. Also recall idea number one: *Living beings are autonomous agents that generate and maintain themselves while enacting their own cognitive domains*. From this idea we gather that through meaning-generation the organism enacts its own cognitive domain in a sort of self-generation process. Thompson's use of 'meaning' here is not one of propositional semantics. Instead, "Something acquires meaning for an organism to the extent that it relates (positively or negatively) to the maintenance norms for the organism's integrity."[62] Given this view, the reference to 'nervous systems' in idea number two is replaceable by 'autopoietic systems'

[62] Thompson, 70. Parenthetic content in original.

in the particular case of biologic cells. In fact, Thompson views the emergence of 'selfhood' in cells as analogous to selfhood in sensorimotor neural networks.[63]

The generation of meaning by either a nervous system or an autopoietic system is a key aspect of the notion of emergence associated with the enactive approach to explaining embodied mind. This meaning-generation is viewed by Thompson as being the enactive emergence of the organism's cognitive world. Therefore, the concept of emergence employed is one of co-emergence, whereby the organism and its cognitive world emerge from each other. This is the circular causality that, in part, constitutes the embodied dynamicism view of mind. The entire system (inclusive of the organism *and* its cognitive environment) is what has emerged – not just the organism. Here is a diagram.

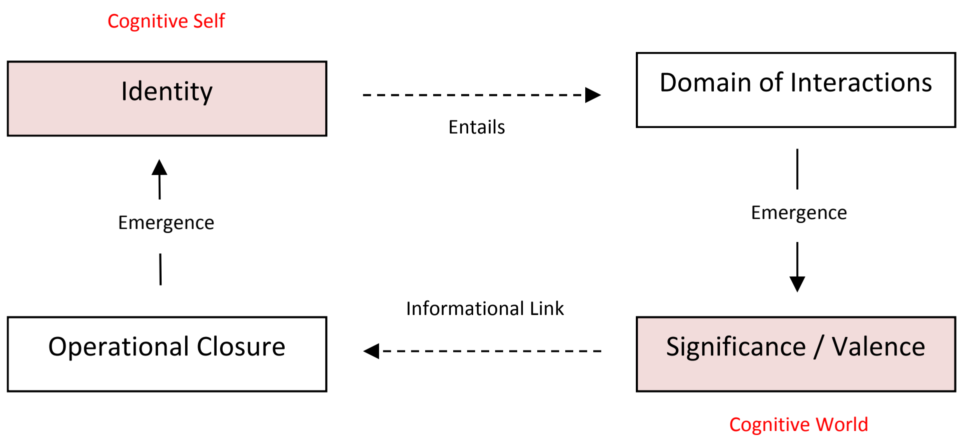

Figure 5. Co-emergence of Selfhood and World[64]

Moving clockwise around Figure 5, identity in the form of a cognitive self emerges from the operational closure of the organism, whereby the organism behaves according to norms that sustain it. That the organism has a self-identity entails there is a domain of

[63] Thompson, 60, lower portion of Figure 3.3.

[64] Thompson, 60, Figure 3.3. Red labels and shading added.

interactions between cognitive-self and other. From this domain emerges a group of interactions having significance (positive or negative) to the sustenance of the organism – i.e. a valence. This valence amounts to the cognitive world of the organism and provides information that influences the operational closure of the system. Under Thompson's view, the co-emergence of selfhood and world is applicable to a number of dynamic systems, including: biologic cells, immune systems, and sensorimotor neuronal networks.[65]

At this stage it is worth focusing on the circularity of Figure 5. Where we begin our analysis in the circle is irrelevant, and the 'up' and 'down' portions are entirely relative. The diagram could be rearranged to place the lower elements on top – provided we reverse the direction of the emergence arrows. Since, under the enactive approach to explaining embodied mind, the co-emergence of world and identity result in 'upward' and 'downward' being relative terms, there is no emergence base. This has implications, Thompson thinks, for Kim's worry over downward causation.

2.5. Groundwork for Thompson's Counter to Kim's Worry

In *Mind in Life*, Thompson devotes a twenty-four page appendix to addressing the problem of downward causation with respect to emergence.[66] He begins by provisionally defining the particular type of emergence he seeks to defend, namely: emergence stemming from nonlinear dynamic systems such that the globally emergent properties have a causal influence on the system's components. His definition is as follows:

---

[65] Thompson, 60, Figure 3.3.

[66] Thompson, 417-441.

"A network, *N*, of interrelated components exhibits an emergent process, *E*, with emergent properties, *P*, if and only if:

1) *E* is a global process that instantiates *P*, and arises from the coupling of *N*'s components and the nonlinear dynamics, *D*, of their local interactions.
2) *E* and *P* have a global-to-local ("downward") determinative influence on the dynamics *D* of the components of *N*.

And possibly:

3) *E* and *P* are not exhaustively determined by the intrinsic properties of the components of *N*, that is, they exhibit "relational holism.""[67]

Thompson's insertion of "And possibly" before the relational holism term of his provisional definition is somewhat confusing in that 'possibly' does not comport with an 'if and only if' definition. It leaves us to wonder whether, if we deleted the relational holism term, the definition would no longer be provisional. He does not say so one way or the other. On the other hand, in his discussion he exemplifies relational holism with instances of quantum entanglement.[68] Then he employs quantum entanglement (referenced as the physics of action at a distance) to deny Kim's claim that synchronic downward causation is incoherent.[69] This would seem to imply he believes relational holism belongs in the definition of emergence – if not he yields territory to Kim's claim of incoherence with respect to the concept of synchronic downward causation.

---

[67] Thompson, 418.

[68] Thompson, 428.

[69] Thompson, 432.

Additionally, Thompson endorses a form of holism when he indicates that his particular notion of emergence is guided by (his view of) contemporary science rather than the 'natural piety' of earlier British emergentism.[70] The particular footnote in which he indicates what he means by 'contemporary science' endorses the Hamiltonian formulation of classical mechanics, whereby energies (not forces) are fundamental. Under the Hamiltonian formulation, Chaos theory describing nonlinear dynamics (the core of Thompson's emergence) views the total energy of a dynamic system (i.e. the Hamiltonian) as nonseparable.[71] In Thompson's words, "According to this proposition, there is a kind of holism proper to complex systems that does not seem compatible with the philosophical doctrine of reductionism. At the same time, this form of holism (i.e. nonseparability of the Hamiltonian) is not incompatible with determinism."[72] Thompson's explicit reference to the Hamiltonian formulation as typifying the contemporary science on which his formulation of emergentism is founded commits him to holism. Additionally, his use of holism to defend against Kim further commits him to holism. Therefore the hedging "And possibly:" before the third term of his if-and-only-if definition of emergence must become simply ‘And:’.

Thompson's definition of emergence qualifies as strong emergence in Kim's sense because: 1) the globally emergent properties are not present in the underlying local components, 2) the globally emergent properties have a deterministic effect on the component interactions, and 3) the holism of the system is not reducible to the

---

[70] Thompson, 437-438 and fn8 reference on 438.

[71] Thompson, see pg. 438 fn8 reference as well as fn8 text on 480. Also see Hamiltonian description on 429.

[72] Thompson, 430.

components of the system. Nonetheless, Thompson distinguishes his view of emergence from many others in that he views the global-to-local *co*-emergence of cognitive individuals and their cognitive environments as being devoid of an absolute up/down distinction. We will consider the significance of this distinguishing feature in more detail shortly.

Thompson divides his attack on Kim in two by first addressing Kim's claim that *synchronic* downward causation is conceptually incoherent, and then addressing Kim's claim that *diachronic* downward causation is either otiose (i.e. functionless) or in violation of the causal closure of the physical domain. I will first consider Thompson's attack on Kim's claim that synchronic downward causation is incoherent.

CHAPTER 3

FAILING TO JUSTIFY DOWNWARD CAUSATION

3.1. The Polemics of Synchronic Downward Causation

3.1.a. A Quantum Non-Locality Claim and Initial Rebuttal

With respect to synchronic downward causation, recall my earlier recital of Kim's "causal-power actuality principle". In essence, this principle holds that causes must already possess their causal efficacy at the time that they exercise that efficacy. In other words, they cannot exercise that efficacy if they do not have it. Therefore it is incoherent to claim that they simultaneously gain it from the very things on which they are having an effect as a result of their causing that effect. With this reasoning, Kim concludes that *synchronic* downward causation is an incoherent concept.

Thompson makes two attempts at justifying synchronic downward causation. I will discuss both. In the first, he attempts to countermand Kim's causal-power actuality principle by pointing to what he sees as simultaneous causation in nature,

> "Simultaneous causation, though counterintuitive to commonsense, is not obviously incoherent. For instance, certain theories in physics postulate instantaneous action at a distance; whether such influences exist, and whether they should qualify as causal, are difficult empirical and conceptual issues."

Instantaneous action at a distance is also referred to as *non-locality*, and it is associated with the phenomenon of quantum entanglement.[73] Thompson's manner of reference to non-locality is puzzling due to the fact that he is using it as a justifying instance of simultaneous causation while also challenging whether it really exists and whether it qualifies as causal. While true that non-locality was initially a difficult phenomenon to empirically confirm, its existence has been verified by the physics community beyond reasonable dispute. I will describe how physicists have come to this conclusion shortly. In doing so, I will also explain why there is a consensus among physicists (though not universal agreement) that non-locality does *not* qualify as a causal phenomenon. If the greater physics community is correct, the notion of non-locality offers no support for simultaneous causation and, therefore, no shield against Kim's criticism of synchronic downward causation. Even if not true, and non-locality could be reasonably construed as causal, it offers no support for the particular *kinds* of macro-level emergent phenomena that Thompson is defending in his book. I will explain.

---

[73] Robert Nadeau and Menas Kafatos, *The Non-Local Universe* (Oxford: Oxford University Press, 1999), 1-4.

But first, it would be appropriate to explain the phenomenon of non-locality – i.e. 'action-at-a-distance'. Non-locality was disparagingly referred to as "voodoo forces" and "spooky interactions" by Albert Einstein.[74] It is a consequence of a foundational principle in quantum physics – *Heisenberg's uncertainty principle*. In general terms, the uncertainty principle holds that the physical features of the microscopic realm divide into two classes, and our knowledge of one class constrains our knowledge of the other.[75] For example, the more precisely we know a particle's position, the less precisely we can know its velocity, and vice versa. This may seem acceptable as a practical matter of measurement in Newtonian physics; however, the principle applies to features like spin, energy, and time as well as position and velocity.

For example, if you are able to detect the x-axis spin of a sub-atomic particle, you are unable to know its y-axis or its z-axis spin. Another example: the more precisely you know the time (i.e. instant) when you measure a particle's energy, the less certainty you have regarding that energy – and vice versa.[76] Keeping in mind that a particle's mass is the bulk of its energy, the uncertainty principle is telling us that the narrower the time interval over which we measure the particle's mass, the less definite the mass. At any absolute instant, a particle's mass is completely indeterminate.

Physicists have wrangled over whether the uncertainty principle is describing an epistemic limitation or, instead, making an ontological proclamation about physical reality. If the latter, then particles do not have a definite mass at any given instant, nor do

---

[74] Bruce Rosenblum and Fred Kuttner, *Quantum Enigma* (Oxford: Oxford University Press, 2006), 137. Also, Brian Greene, *The Fabric of the Cosmos* (New York: Random House, 2004), 80.

[75] Greene, *The Fabric...* , 95-96.

[76] Brian Greene, *The Elegant Universe,* (New York: W. W. Norton & Company, 1999) 115-116.

they have definite spins on all three axes, nor do they have both location and velocity at the same time. Einstein waged war on the ontologic interpretation of quantum physics and the uncertainty principle beginning in 1935 with a paper jointly authored by him, Boris Podolsky and Nathan Rosen – now famously referred to as the EPR paper.[77] He continued that fight to his death in 1955, and was eventually proven to be wrong via a series of experiments conducted in the 1970s and early 1980s.[78]

An often-cited proof that Einstein was wrong came from the work of a French physicist, Alain Aspect, and his colleagues.[79] The experiments involve exciting calcium atoms and then allowing them to return to their normal state. When an excited calcium atom returns to its normal state, it emits two photons traveling in opposite directions with *correlated* spins, whereby if one is spinning clockwise, the other will be spinning counterclockwise.[80] In the Aspect experiments two detectors for measuring photon spin are placed at equal distance on opposite sides of the calcium atoms.[81] When photon spins are measured at the left detector, random spin results are found. When measured at the right detector, random results are also found; however the *random* results on the right

[77] Greene, *The Fabric ...,* 99-102.

[78] Greene, *The Fabric ...,* 112-115.

[79] Greene, *The Fabric ...,* 113. Nadeau & Kafatos, 78. RosenBlum & Kuttner, 149. John Gribbin, *Schrödinger's Kittens and the Search for Reality* (Boston: Little, Brown and Company, 1995), 23. Aspect was trying to discover whether a theorem developed by physicist John Bell (often referred to as 'Bell's inequality') was true. That theorem provided a framework for testing whether or not Einstein was correct in asserting reality is local. If Bell's inequality is violated, the universe is nonlocal. Aspect's experiments demonstrated violation of Bell's inequality. Nadeau and Kafatos, 69-70.

[80] Technically the spins are 'anti-correlated' meaning always exactly opposite; however, the literature often uses 'correlated' as a convenience. Greene, *The Fabric ...,* 166.

[81] Brian Greene describes the Aspect experiment as measuring photon spin while other physicists describe measuring the direction of polarization. When thinking of light as a wave physicists use polarity and when thinking of light as particles (i.e. photons) they use spin. The two notions are equivalent in that polarization amounts to the average spin directions of the photons detected. Greene, *The Fabric ...,* 432.

detector precisely correlate to the *random* results of the left. Whenever the right is found to be spinning clockwise, the left is found to be spinning counterclockwise – and vice versa. In order to carry out these observations, one needs to simultaneously collect a list of measurements from both detectors and, after the fact, compare the lists. Due to the correlation of the two lists, the particles are said to be entangled. Yet the lists have no discernible pattern when studied individually. They are completely random. Physicist Brian Greene remarks on the randomness of the photon pair spins and offers an analogy:

> "They are like a pair of magical dice, one thrown in Atlantic City and the other in Las Vegas, each of which *randomly* comes up one number or another, yet the two of which somehow manage always to agree. Entangled particles act similarly, except they require no magic. *Entangled particles, even though spatially separate, do not operate autonomously.*"[82] [Emphasis in the original.]

Quantum entanglement of particles has been demonstrated to distances of eleven (11) kilometers, and there is no reason to believe that entanglement would not be true across billions of light years.[83] Given that entangled particles show random but correlated spins, we can conclude their ontological status is that of a kind of whole across space. If not for the randomness we could argue that the photons merely leave the calcium atom with correlated spins and behave as individuals thereafter. However, the randomness of the measurements at a later time (and further distance) precludes the particles from being fully autonomous individuals since their random spin results are correlated no matter when or how far away we measure them. If they were behaving

---

[82] Greene, *The Fabric ...*, 114.

[83] Greene, *The Fabric ...*, 115.

autonomously as they moved apart, there would be no correlation between their randomness. Physicists take this as confirmation that quantum uncertainty reflects the ontological status of our world and is not merely an epistemological limitation. Thompson and others take it to show that holism is true – i.e. that the emergent features of the whole have a synchronic *causal* effect on the components.

From the Aspect experiments and others like it, we are safe in following the lead of contemporary physicists and concluding that action at a distance *does* exist. Nonetheless, there is a not sufficient reason to conclude that action at a distance qualifies as *causal*. This is a subtle, though important, distinction when considering quantum entanglement as a possible justification for emergence involving synchronic downward causation.

A pair of quantum entangled particles will become disentangled once either of them is disturbed by, for example, measuring the spin. The process of quantum states becoming disentangled is referred to as "decoherence".[84] When a pair of entangled photons decoheres, the disentangling happens instantaneously – regardless of their separation distance.[85] If I measure the spin of the left photon as clockwise, the instant in which that photon is determined to be clockwise the right photon is counter-clockwise. The pair then becomes two distinct autonomous individuals, whereby – going forward in time – their spins will no longer be correlated. This should give us pause. Einstein's well-verified theory of special relativity tells us that *nothing* can travel faster than light.

---

[84] John Gribbin, *Q is for Quantum, An Encyclopedia of Particle Physics* (New York: The Free Press, Simon & Schuster, 1998) 101.

[85] Greene, *The Fabric ...*, 115-116.

How can the right photon *instantaneously* gain the information about which random spin the left is undergoing to ensure that the right correlates exactly opposite?

The answer is: There is no information for the right to gain about the left since they are an entangled whole. They can exist in this state of a whole (spread across space) because, in the majority view of physicists, particles in isolation have a probabilistic existence, whereby their location and attributes are expressed as probabilities rather than concrete certainties. As Greene puts it:

> "Particle properties, in this majority view, *come into being* when measurements force them to ... When they are not being observed or interacting with the environment, particle properties have a nebulous, fuzzy existence characterized solely by a probability that one or another potentiality might be realized."[86] [Emphasis added.]

At the moment that a measurement of an entangled pair causes the whole to decohere into two now-autonomous photons, nothing travels between them – no particles, no waves, no energy, no signal and no information. We can rightfully claim that no information passes between them because 'they' are really and 'it' – i.e. a single physical entity. One way to verify this is to ask whether we could use quantum entanglement to transmit an instantaneous message across a great distance. The answer is no, and here is a thought experiment to show why.

You and I initially meet in the laboratory next to the calcium atom exciter while it spews entangled pairs in opposite directions. We then set our portable atomic clocks to

[86] Greene, *The Fabric* ..., 121.

precisely the same time, and we move in opposite directions – carrying our clocks and photon spin detectors – to exactly five kilometers away from the exciter. We agree, in advance, that at a precise moment we will interrupt the beam of photons to measure spin. If you measure the spin of the photon on your side at exactly that same moment and verify that it has clockwise spin, you will know that the photon I am measuring at that precise moment has counter-clockwise spin. Have we transmitted information? No. The spin measured is random. I could, of course, alter the spin of my photon; however, the moment I jostle it to do so, it loses its correlation to its paired other. What I do to mine has no influence on the future spin of its prior-paired other, since they are no longer unified under entanglement.

You might object that you would at least know I am now influencing spins (at least on my side), because the photons you are measuring are no longer in coherence with a paired other. Not so. You would have no way of knowing whether it is your own measurement or mine that decohered the photon you now measure. The best that either of us can do is keep a list of the spins we measure and compare them via some (non-instantaneous) *light-speed or slower-than-light-speed* information transmission process, and only *that* comparison would be an information transmission qualifying as a causal event. The actual decoherence of the whole into two autonomous photons may well be a causal event; however, the fact that one photon turns up with clockwise spin is not the cause of the other turning up counter clockwise. Brian Greene describes the results of an analogous thought experiment in this way:

> "Thus, there is no message, there is no hidden code, there is no *information* whatsoever in either of these two random lists. The only interesting thing about the two lists is that they are identical–but that can't be discerned until the two lists are brought together and compared by some conventional, slower-than-light means... The standard argument thus concludes that although measuring the spin of one photon appears instantaneously to affect the other, no *information* is transmitted from one to the other, and the speed limit of special relativity remains in force. Physicists say that the spin results are correlated–since the lists are identical–but do not stand in a traditional *cause-and-effect* relationship because nothing travels between the two distant locations."[87] [Emphasis added.]

When I studied this passage the first time through I found myself squinting with suspicion. If we take 'affect' in its Oxford dictionary sense of "have an effect on"[88], how is it that one photon can instantaneously have an effect on the other without there being cause involved? Well, notice that Greene uses the term 'appears', and he does so because he realizes the one photon is not really having an effect on the other. They are a sort of whole. On the other hand, the thought experiment clearly indicates that, when one photon becomes disentangled via measurement so as to yield a clockwise spin, the other is forced to be counterclockwise. It seems clear that the *information* about spin is being passed from one photon to the other despite Greene's claim that they are, in some sense, a physical whole. It seems the measured photon *causes* the other to adopt opposite spin.

---

[87] Greene, *The Fabric ...*, 117.

[88] *The New Oxford American Dictionary,* second edition, s.v. "affect" 25.

3.1.b. Clarifying What We Mean by 'Information' and How to Apply the Notion to Quantum Non-Locality

Here might be a good place to clarify how we ought to think of the terms 'information' and 'cause' as we apply them to this discussion. Information has a long standing colloquial sense as well as a more recent technical definition. Physicist Hans Christian von Baeyer likes to communicate the colloquial sense of information using the term "in-formation".[89] The idea is, "Information refers to molding or shaping a formless heap – imposing a form onto something. So the question of its meaning reverts to the more fundamental one: What is *form*?"[90] [Emphasis in original.] In contrast, "Randomness is defined as lack of form or pattern, so the adjective 'random' means 'formless, patternless, and therefore unpredictable'."[91]

Here we gain some insight into Greene's rationale for claiming that no information is in either of the random lists of spins collected by our two detectors, and why no information is transmitted from one photon to the other upon decoherence. We can assume that Greene is using the colloquial sense of information-as-pattern, and lack of information as randomness. Then we imagine 'entangled whole' not as a pair of photons with opposite spins, but as a single entity with nebulous, probabilistic attributes in terms of location and spin. If we perturb the whole via a measurement, its probabilistic attributes become determinate certainties. It goes from being a whole to two halves – i.e. two photons. By analogy, just as a whole becoming divided in space can be said to have

[89] Hans Christian von Baeyer, *Information, The New Language of Science* (Cambridge: Harvard University Press, 2003), 18-27.

[90] von Baeyer, 20.

[91] von Baeyer, 100.

a right half and a left half, an entangled particle pair whole becoming divided in space will have a clockwise spin and a counterclockwise spin. Prior to the divide (i.e. prior to decoherence), the spin is an indeterminate probabilistic attribute; and the laws of nature at the quantum level are such that it can only be divided in symmetric clockwise-counterclockwise halves. Whether the right turns up clockwise and the left counterclockwise or vice versa is an entirely random matter in the Aspect experiment. The probability function describing the likelihood of clockwise versus counter clockwise (known as a probability wave function) is spread throughout the entire universe.[92] The Aspect experiment is telling us that, when that wave goes from being a probability to an actual certainty, it becomes certain everywhere at once.

If the majority view of quantum physics is true, then synchronic causation is not supported by the action-at-a-distance of non-locality, and Kim's claim about the incoherence of downward synchronic causation remains a troubling threat to strong emergence. One possible avenue of attack on the physics community's consensus is to claim the *colloquial* sense of information is misplaced when used as an explanation of scientific test results. Science requires *technical* definitions and the technical definition of information is agnostic with respect to randomness in content. To see how this argument would be prosecuted, consider first the technical definition of information.

Claude Shannon, the founder of modern information theory, is credited with the technical definition of information. Von Baeyer describes 'Shannon information' as follows, "*To find the information content of any message, translate the message into the*

---

[92] Greene, *The Fabric ...*, 90. Gribbin, *Q is for Quantum,* 293-4. Rosenblum and Kuttner, 102-3.

binary code of computer and count the digits of the resulting string of zeroes and ones." [Italics in original.][93]

Von Baeyer notes that Shannon information is not really about content as much as quantity. The Shannon number is a measure of the *amount* of information in a message – whether meaningful or gibberish (i.e. random). Though we might prefer not to think of gibberish as information, the concept of having a reliable standard to measure the amount of information in a given transmission is very helpful in, for example, designing telephone transmission systems, or satellite communications systems, or software data compression algorithms. The significance and usefulness of the Shannon number comes, in part, from Shannon's mathematical proof that binary code uses the smallest possible amount of resources in terms of both memory and transmission bandwidth. Therefore, the Shannon number is a genuine measure of the maximum amount of information that we could glean from a message. In addition to serving as a measurement standard for the design of information processing systems, it can help us analyze whether we have enough information to make a decision or hold a belief.

To see how this works, consider my Chapter 1 lament regarding the fuzzy border distinguishing B-weak emergent phenomena from phenomena that are borderline multiply realizable – and therefore perhaps not B-weak emergent. I gave the imaginary example of alt-life under which there existed a grand total of only two 'hopper guns'. Since the life grid extends infinitely in two dimensions, one could never try every possible initial combination of life forms to discover whether the two hopper guns were,

[93] von Baeyer, 28.

in fact, the only two. Now consider Conway's Game of Life with glider guns instead of hopper guns. There are many more than two glider guns, though we don't know how many. Suppose, for simplicity, there are sixteen – in fact there are many more.[94] And suppose that we wrote two computer programs to randomly search initial life combinations – one running alt-life rules looking for hopper guns, and the other running Conway-life rules looking for glider guns. We launch both programs and run them for one billion different initial configurations each. At the end, each will print a message indicating how many guns were found. The alt-life computer prints '2' and the Conway-life computer prints '16'. Converting their messages to binary and finding the Shannon information content yields the following.

Table 1: Shannon Information

Decimal	Binary	Shannon	
1	1	1	
2	10	2	Alt-Life Hopper Guns
3	11	2	
4	100	3	
5	101	3	
6	110	3	
7	111	3	
8	1000	4	
9	1001	4	
10	1010	4	
11	1011	4	
12	1100	4	
13	1101	4	
14	1110	4	
15	1111	4	
16	10000	5	Conway Life Glider Guns
17	10001	5	
18	10010	5	

[94] For fun you may want to web-search: 'how many Game of Life glider guns exist?'

The wide disparity in the Shannon information content of the two results confirms my prior conclusion that we have good reason to worry whether the hopper gun phenomenon in the alt-life world is genuine B-weak emergence. In contrast, the relative amount of information available to us in making our claim about B-weak emergence with respect to glider guns is 2 ½ times greater than that of hopper guns.

This same procedure can be related back to Brian Greene's conclusion about the correlated-but-random lists of photon spins in the quantum entanglement experiments – though perhaps not in Greene's favor. On the positive side, a long list of correlating spin results will have a high Shannon number indicating we have enough information to conclude that the photons were, in fact, entangled. Physicists have compiled such lists in laboratory experiments to the point where serious challenges to the phenomenon of quantum entanglement are many years in the past.[95] On the other hand, champions of emergence via synchronic downward causation can point to the Shannon definition of information to discount the randomness of the spin result lists. Greene's assertion that no information is transmitted during decoherence of quantum entangled pairs is based on the lists being random. Under the parochial definition of information, random strings have no meaningful content and therefore carry no information. However, under the technical definition of information, the Shannon number can be used to assert the presence of information *regardless* of the meaningfulness of its content. That the left list and the right list are each random is irrelevant – from Shannon's perspective. Randomness *is* information, and the correlation of the randomness implies information has been

[95] Greene, *The Fabric ...*, 115.

transmitted. Therefore, Greene's assertion that no information has been transmitted is questionable – *if* we accept that randomness contains information.

This leaves a narrow opening for defenders of synchronic causation. They may assert: *Even if we accept a description of the entangled pair as having nebulous probabilistic attributes, the fact that we can untangle the whole into a pair by measuring either the right photon or the left photon, implies the measured photon* <u>*causes*</u> *its paired partner to adopt an opposite spin*.

3.1.c. Clarifying What We Mean by 'Cause' and How to Apply the Notion to Quantum Non-Locality

Before we can fully reply to this assertion, we need to consider use of the term 'cause'. Aristotle famously explored four senses of 'cause' in *Book II* of his *Physics*.[96] I will use the traditional example of a statue to distinguish those four senses:

1. Material Cause – "that out of which" the statue is made, e.g. marble;
2. Formal Cause – "the form", "the account of what-it-is-to-be", e.g. the statue is formed so as to be like David the slayer of Goliath;
3. Efficient Cause – "the primary source of the change or rest", e.g. Michelangelo is the primary source of the change from block to statue;
4. Final Cause – "the end, that for the sake of which a thing was done", e.g. evoke civic pride in the Republic of Florence and prestige of the patrons.[97]

[96] Aristotle, *Physics, Book II,* 195a 15-25, *Aristotle Introductory Readings*, translated by Terence Irwin and Gail Fine (Indianapolis: Hackett 1996).

[97] Andrea Falcon, "Aristotle on Causality", *The Stanford Encyclopedia of Philosophy (Fall 2011 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2011/entries/aristotle-causality/>.

When an advocate of synchronic causation asserts that measuring one photon causes its entangled partner to adopt an opposite spin, we can safely assume that the sense of cause being used is efficient cause – i.e. the primary source of change. By elimination, the advocate is not making a statement about the stuff that composes the photon pair, nor about the behavior that comprises spin; therefore, the sense is not material cause. Similarly, the advocate is not making a claim about the form of the photon pair or whether its form or spins are supposed to represent or be analogous to something; therefore, the sense in use is not that of formal cause or an 'account of what it is to be' an entangled pair becoming disentangled. Finally, the advocate is not making a teleological claim about the purpose of the photon pair or its particular spins or behaviors or of the experimenter's purpose; therefore the sense in use is not that of final cause.

The advocate *is* claiming that the primary source of the photon's change from a nebulous state of indeterminate quantum probabilities to a determinate state of, for example, clockwise spin, is the counterclockwise spin of its paired partner. In other words, the efficient cause of the clockwise spin is the counterclockwise spin of the other photon. This sense of cause comports with what Ned Hall refers to as 'productive' cause and what Jaegwon Kim likes to call 'productive/generative' cause.[98] Causes generate and produce their effects. I view this 'productive/generative' sense of cause as approximately equivalent to asserting causes are the primary source of change – and therefore approximately equivalent to efficient cause. This sense of cause will be particularly relevant to our exploration of emergence and mental causation, and my use of 'cause'

[98] Kim, 255-257.

with respect to mental causation in the upcoming subsections of this paper will the productive/generative sense.

In regard to mental causation, Kim has explored generative cause in contrast to cause as nomological relationship (i.e. law-like regularity between cause and effect). He has also contrasted it with counterfactual dependency (i.e. the effect would not have happened had not the object or phenomenon been there). Counterfactual dependencies often play a role in the outcome of an event without being the productive/generative cause. For example, when a lead bullet is fired from a gun and strikes a concrete wall it becomes deformed. Had the wall not been there, the bullet would not have been deformed. We can only claim the wall 'caused' the deformation in the true counterfactual sense of cause, but not in the productive/generative sense. The exploding gun powder was the productive/generative cause for the bullet's deformation. Note that both the exploding gun powder and the wall qualify as true counterfactuals; however, only the exploding gun powder qualifies as the productive/generative cause.

Kim's interest in the study of cause is to determine which sense of cause we are using when we ask whether mental causation is a coherent concept or, instead, reduces to neural/physical realizers. In his view, the productive/generative sense of cause is the only sense that captures our meaning when we ask whether mental causation is 'real'. We are asking whether mind generates and produces our actions.[99] I agree with Kim's assessment. The productive/generative sense of cause is precisely what scientists are worried about when they insist that non-physical entities cannot produce energetic

[99] Kim, 262.

changes in physical entities. It is the sense of cause that derives from the well proven natural law of conservation of energy, and it is the sense of cause referenced when scientists assert the universe is causally closured. Furthermore, this is the sense in which Thompson uses 'cause' when he claims the global emergent structure of dynamic complex systems bring into being new causal powers that produce local effects. Throughout the remainder of this work, I will use 'cause' in Kim and Hall's productive/generative sense which is, I believe, roughly equivalent to Aristotle's efficient cause. As I present arguments aimed at making progress on the issue of mental causation, it will be important to keep in mind that I am not using 'cause' in the true counterfactual sense, because the truth status of many assertions about mental causation will often rest on whether the thinker has in mind productive/generative cause or passive true counterfactual cause.

3.1.d. Finishing the Quantum Non-Locality Rebuttal

Now, reconsider our quantum entangled photon pair. The advocate for synchronic downward causation is claiming the *efficient* cause of the clockwise spin for one photon in a recently disentangled pair is the counterclockwise spin of the other photon. The spin of one photon has generated/produced the spin of the other. This claim is incorrect. Recall that the empirical evidence of the Aspect experiment shows a given pair of entangled photons will adopt their correlated, clockwise-counterclockwise spins simultaneously. The primary source of the change from indeterminate probability to actual spins is the measurement of the entangled whole – *regardless of whether that measurement is taken on the left or the right*. The reason why, for example, the left spin turns up clockwise while the right simultaneously turns up counter clockwise is that the this particular probability set became actual. There was an equal probability that the reverse could have become actual, and there was *no* probability that the spins could have turned up the same – any more than a tossed coin can turn up both heads and tails. Sticking with the coin analogy, we do not claim that the efficient cause of a tossed coin turning up heads is the fact that the tails side is down. In the same way, we cannot claim that the efficient cause of an entangled pair turning up clockwise on the left is due to the fact that it turned up counterclockwise on the right. Therefore, even if we adopt the notion that Aspect's experimental results are Shannon information, the follow-on claim that even random data amounts to information does nothing to support the claim that the right photon caused the spin of the left or vice versa. The efficient cause of *both* spins was the disturbance of the entangled whole via measurement. Both spin outcomes are due to a particular probability set becoming actual.

3.1.e. The Claim for Rayleigh Bénard Convection Cells

I began the discussion of synchronic downward causation by indicating Thompson offers two counter arguments to Kim's claim that synchronic downward causation is an incoherent concept. The first Thompson argument was an attempt to use quantum non-locality as an example of simultaneous causation. I believe I have shown that attempt fails. Though my reasoning against the argument from quantum non-locality was somewhat laborious, I used it as an opportunity to explore and clarify two concepts that will prove useful in the upcoming discussion about emergence and mental causation – i.e. 'information' and 'cause'.

Thompson's second attempt to justify synchronic downward causation is an explicit attack on Kim's 'causal-power actuality principle'. In that attack, Thompson cites the macro-level phenomenon of Rayleigh-Bénard convection rolls as direct evidence against Kim's principle.[100] This attempt against Kim also fails for reasons that will be considerably easier to convey than wading into quantum physics. First, I will present Thompson's argument.

Recall again Kim's causal-power actuality principle. In coarse terms it holds that a cause must come before its own effect. Even if the interval by which it precedes its effects is vanishingly small, the sequence of cause first and effect next is inviolable. Against Kim, Thompson asserts that the phenomenon of global-to-local structuring influence is evidence against the causal-power actuality principle. He cites Rayleigh-Bénard convection rolls as examples.[101] These rolls (also called convection cells) are orderly patterns of convective motion that emerge out of the local dynamics of fluid motion. They appear, for example, when an otherwise undisturbed viscous fluid is heated from below, and they take the form of an array of adjacent hexagonal columns.[102] Thompson cites philosopher Robert C. Bishop in asserting that the governing property causing the fluid to transition from a non-patterned state to hexagonal cells is an instance of synchronic, global-to-local structuring. In other words, at some particular instant, all of the molecules in the affected area are simultaneously caused to begin behaving according to a new global order.

---

[100] Thompson, 432-433.

[101] Thompson, 433.

[102] McGraw-Hill Dictionary of Scientific and Technical Terms, Sixth Edition, (New York: McGaw-Hill, 2003), 1753.

3.1.f. Rebuttal of the Rayleigh Bénard Claim

This claim should give us pause. In order for it to be true, all of the molecules in both the horizontal and vertical directions must instantaneously switch behavior. There can be no wave of behavior change propagating in either a horizontal or a vertical direction. Otherwise, the global-to-local structuring would not be truly synchronic. Here we run into the very same concerns that were raised with quantum entanglement; however, in this case we are not dealing with quantum-level particles like photons isolated in vacuum chambers and shielded from observation so as to protect their state of quantum uncertainty from decoherence. We are instead dealing with macro-level molecules in fluids. Those molecules are in intimate contact with each other and the world around. They are not in quantum isolation. Nor are they entangled.

For the macro-level molecules throughout the affected area to *simultaneously* begin behaving according to a new structural order due to genuine synchronous causation, the cause of that behavior must propagate through the liquid faster than light – otherwise Thompson proponents cannot claim *synchronous* causation. However, Einstein's special relativity explicitly forbids anything – including matter, energy or information – from being transmitted with a velocity faster than light relative to any other object.[103] As physicist John Gribbin points out: "The special theory of relativity is not just some crazy hypothesis, but passes Newton's experimental test – it 'explains the properties of things' *and* 'furnishes experiments' which can be used to test (successfully) those

[103] Greene, *The Fabric ...*, 116-117.

explanations."[104] Scientists have had more than 100 years to show that special relativity is 'just some crazy hypothesis', and many have tried.[105] Its success at withstanding scrutiny has been without exception. Therefore, there is little reason to believe Bishop's assertion that the causal transition from non-structured to structured behavior propagates throughout the three dimensional volume of fluids instantaneously. If the causal transition is not propagated instantaneously, then Rayleigh-Bénard convection cells are *not* examples of genuine synchronic global-to-local causation. They might exemplify global-to-local causation in the form of structured behavior; however, the transition to that behavior is diachronic according to special relativity. Any test to prove that the transition to structured behavior is genuinely synchronous, is an attempt to prove that special relativity does not apply in this case. It has the same probability of succeeding as have all other (failed) attempts to discredit special relativity.

The argument from special relativity is more than enough to demonstrate that genuine synchronic global-to-local causation in the form of Rayleigh-Bénard convection cells is fantastically improbable. Nonetheless, we can reason to a similar conclusion from an entirely different direction. Here is how.

The reason why these Rayleigh-Bénard convection cells are also referred to as convection rolls is that the fluid within the individual cells is rolling in a vertical loop due to thermal convection. Warmer fluid rises, cools, and then descends. The only way for this convection roll to occur is for there to be a temperature differential driving the thermal convection. In laboratory experiments this is typically caused by heating at the

---

[104] Gribbin, *Schrödinger's Kittens and the Search for Reality*, 79.

[105] Einstein came up with special relativity in 1905. Greene, *The Fabric ...*, 9.

bottom of the fluid container such that there is a temperature gradient from bottom to top. It can also be caused by merely insulating the bottom side of a very hot viscous liquid pool and leaving the top exposed to cooler air. For example, under certain circumstances, hot lava that stagnates into a large still pool will form convection cells that, over time, cool and solidify into massive basalt rock pillars with imperfect hexagonal and pentagonal shapes. I have marveled at these sorts of geometric rock pillars during my visits to Sheepeater Cliffs at Yellowstone National Park.[106]

We can perform a thought experiment to see how the transition to Rayleigh-Bénard global order cannot be synchronic. Consider an experiment in which a flat-bottomed, smooth dish of oil is heated uniformly and gradually from beneath. As its temperature is increased there will come a point when the fluid seems to synchronously transition into a hexagonal structure of behavior as viewed from above; however, the synchrony is an illusion. Throughout the experiment there was a thermal gradient from the bottom of the dish to the surface of the liquid. As the mass of liquid was heated to approach the point of global-to-local structuring, the threshold temperature necessary for that structuring had to occur *first* at the bottom of the dish and propagate (sub-light-speed) to the surface – at which point the formation of hexagonal columns became complete. The efficient cause of the transition from random, non-structured thermal motion to structured thermal motion was a vertically-traveling wave of threshold temperature. This vertical

---

[106] Sheepeater Cliff was mistakenly named after a tribe of Native Americans thought to live nearby. For information about Yellowstone's Sheepeater Cliff and its hexagonal basalt columns go to the U.S. National Park Service link: http://www.nps.gov/yell/photosmultimedia/0006sheepeater-cliff-iy.htm, Visited 10/20/12.

propagation was diachronic. Therefore, Rayleigh-Bénard convection cells are not emblematic of synchronic global-to-local (i.e. downward) causation.[107]

Given the failure of both of Thompson's examples of synchronic downward causation – i.e. quantum entanglement and Rayleigh-Bénard cells – Kim's assertion that synchronic downward causation is an incoherent concept still stands. Causes must come before their own effects. Our field of debate is therefore limited to diachronic downward causation.

3.2. The Polemics of Diachronic Downward Causation

3.2.a. Overview Kim's Worry and Thompson's Response

So what about Kim's worry over diachronic downward causation and Thompson's rebuttal? Recall my 'Figure 3' example depiction of mental causation (pg. 30). If we accept that the mental state of experiencing insult can cause anger, and we also accept the empirical evidence suggesting each mental state is caused by an underlying physical neural state, then the state of anger is overdetermined. It has two causes, and each cause seems sufficient. This is a problem for the champions of strong emergence – including those who support Thompson's notion of globally emergent phenomena in dynamic systems. Emergent phenomena arising at the global level of a dynamic system are characterized by *new and distinctive* causal powers that influence the behavior of their basal components. Yet, with respect to the overdetermination scenario and mental

[107] The same conclusion will obtain in the case of cooling lava in still pools if one imagines the time immediately after the lava goes from flowing to still. While the lava was moving there was greater thermal homogeneity due to mixing. When the lava becomes still, a thermal gradient begins to form starting from the cooler surface and moving downward into the hot fluid – enabling Rayleigh-Bénard behavior along the way. Therefore the efficient cause of the transition from nonstructured behavior to structured behavior is a vertically traveling thermal gradient wave. This vertical propagation is diachronic not synchronic.

causation, those claimed powers are less than distinctive in that they are completely unnecessary. The causal work is accounted for at the local level of neural activity, and therefore epiphenomenalism looms.

The overdetermination worry is even more troubling when we try to explain human actions triggered by phenomenal experience. Suppose the insult in our example was received during a phone conversation, and the ensuing anger caused the hearer to hang up the phone. Were the nerve impulses that led to the muscle movement caused by the mental experience of anger or by the neural correlates beneath that anger? If emergent minds have genuine causal powers, those powers would seem otiose (to use Kim's characterization) given that our empirical sciences can trace a chain of neural causes leading back from the muscle movement to the brain states beneath the experience of anger and further upstream to the brain states beneath the insult.

Thompson views Kim's argument against diachronic downward causation as inflexible, part/whole reductionism that improperly sets apart phenomenal consciousness as being resistant to physical reductive analysis. He claims Kim's view is, at base, a Cartesian dualism conception of nature.[108] In contrast, the greater portion of Thompson's work in *Mind In Life* is devoted to overcoming dualism by advancing his enactive approach of explaining embodied mind. Cognition is seen as involving a living being in the form of a dynamic, self-maintaining system that generates its own cognitive world by creating meaning through recurrent sensorimotor interaction. The 'meaning generation' associated with the incipient stages of cognition is not the sort of meaning that we

[108] Thompson, 438.

typically think of in terms of representational semantics. Instead, something is meaningful to an organism to the extent that it relates favorably or unfavorably to the norms of maintenance required to preserve the organism's operational integrity.

By shedding the notion of representation processing as a prerequisite for cognition, Thompson moves away from Cognitivism's metaphor of mind as software. He also moves away from Connectionism's view of mind-as-neural-network. Embodying cognition as a kind of know-how at the lowest levels of life helps us to avoid the confounding Cartesian dualism that has occupied so many philosophers for the past three hundred and fifty years or so. It also provides Thompson a platform from which he can attempt to undermine Kim's worry of overdetermination. How so? Because emergence, under the enactive approach to embodied mind, has no definitive up or down. Minds co-emerge with their environments, and wholes co-emerge with their parts.

Refer again to Figure 5 'Co-emergence of Selfhood and World' (pg. 43). This figure is a slightly modified version of Thompson's own.[109] It displays circular causality, whereby up and down are irrelevant as 'self' emerges. Let's review again how that emergence takes place, this time paying attention to what the 'self' comprises.

The self-identity of the organism emerges from the operational closure of the organism. This self-identity entails that a domain of interaction must exist between that self and its environment. From that domain of interaction, the world of the organism (i.e. its significance valence) emerges. This significance valence, in turn, serves as an information source for the operational closure of the organism. In other words, the self-

---

[109] Thompson, 60.

maintenance of the organism – its operational closure – is driven by its ability to respond appropriately to stimuli that it finds meaningful in relation to its norms of self-maintenance. Notice that, under this view, the 'world' of the organism is not the entire physical world that the organism comes in contact with. It is only the portion of that physical world having significance relative to the norms of maintenance required for the organism to thrive. Though the entire physical and microphysical world exists before the organism, its cognitive world emerges from the organism while the organism emerges from that cognitive world. Under this view, the organism is not constituted by its atoms and molecules since those are in large part – if not completely – replaced over its lifetime. Instead, the organism is constituted by the *process* portion of the closed loop system that operates to maintain the viability and self-identity of the organism. Under Thompson's view, the emergence does not occur upward from micro-constituents; but, instead, it co-occurs between the processes of organism and significance environment. The self is not a grouping of atoms or molecules. It is a stabilized process existing over time.

Thompson applies this same notion of dynamic, self-reflexive co-emergence to the brain in higher order organisms such as humans. In addition to seeing a closed loop of co-emergence operating between organism and environment, he sees a closed loop of co-emergence operating at a higher level within brains. Following the work of biologist philosophers H. R. Maturana and F. J. Varela, Thompson understands the brain operationally as a dynamic network of processes displaying part/whole as well as whole/part co-emergent behavior. Thus, "One might as well say that the components (local neuronal activities) emerge from the whole as much as the whole (dynamic patterns of large-scale integration) emerges from the components."[110] [Parenthetical content in original.] In other words, our mental/experiential lives do not emerge from neurons. They emerge from processes of local neuronal activities, and those same local neuronal activities (i.e. the parts) emerge from the global processes that compose our conscious minds (i.e. the emerged whole).

Given this particular view of emergence, how does Thompson defend against Kim's worry over diachronic downward causation? He begins by noting that his context of reference is complex systems, and, as such, the phrase "emergent downward causation" applies only to "... the specific kind of reflexive global-to-local influence that happens in a system that has dynamic global coherence in and through collective self-organization."[111] After describing Kim's downward causation worry, Thompson complains that Kim's way of thinking fails to acknowledge the causal role that the macro-level organization of a complex system plays on its micro-level interactions. He asserts the macro-level organization is not simply epiphenomenal, and he offers two reasons why not. First, it is multiply realizable, and second, true counterfactuals associated with macro-level systems can be different than true counterfactuals associated with their micro-level processes.[112] I will explain and question each.

3.2.b. Multiple Realizability and Claims of Irreducible Chemistry

Regarding the first, Thompson uses biologic cells to exemplify his 'multiply-realizable' argument. Cells continually regenerate themselves – as do all living things.

---

[110] Thompson, 423.

[111] Thompson, 434.

[112] Thompson, 436-437.

The invariant thing is the organizational structure associated with the individual – not the micro-physical matter. Since its constituent parts are exchanged over time, the cell is multiply realizable across time. Therefore, Thompson argues, the global organizational structure that emerges from the constituent parts is far from merely epiphenomenal. In fact, that emergent structure is what constitutes the cell as an individual organism.

This argument appears compelling; however, it loses its pull when we ask whether the global behavior of a cell is reducible to the behavior of its constituent parts using the laws of chemistry and physics. Keep in mind that we are not examining B-weak emergence, whereby: a) Fully reducible behaviors have the emergence-signaling trait of being multiply realizable *without any claim of new causal powers* coming into existence, and b) 'multiply realizable' refers to instantiations in separate individuals rather than a single individual across time. Here we are, instead, examining strong emergence, whereby emergent phenomena are claimed to bring with them new causal powers and those powers are not reducible to their basal conditions. They are said to affect the constituent parts from which they have emerged. Advocates of strong emergence must worry about reducibility. If the global behavior of biologic cells reduces to chemistry and physics – whether multiply realizable or not – the purported new causal powers of strong emergence have no work to do, and the emergent global structure is therefore epiphenomenal. In other words, the structure itself is causing nothing to happen that is not already fully explained by the properties and behavior of the underlying constituents.

Thompson anticipates this line of attack and asserts that inter-theoretic unity between chemistry and physics is "nowhere in sight".[113] In other words, the laws of chemistry that underlie and determine so much of biology are not fully explainable by or reducible to the laws of physics. If the purported gaps between physics and chemistry are an intrinsic feature of nature (as opposed to an epistemic gap), then there is both room and need for new causal powers as posited by strong emergence.

Thompson is not alone in his claim against inter-theoretic unity. In fact, the fairly recent discipline of Philosophy of Chemistry trends heavily against unity.[114] Nonetheless, philosophers of chemistry are not in the majority among deep thinkers on this issue, and the holdouts arguing in favor of reducibility are more likely to be physicists and chemists. For example, chemist Henry Bent cites the union of chemical and physical theory in his physics encyclopedia entry on chemical bonding. Likewise, physicist John Gribbin asserts that chemistry is a branch of physics, and physicist Leonard Susskind agrees in a separate publication.[115] Counter to their claims, Thompson

---

[113] Thompson, 440.

[114] Michael Weisberg, Paul Needham, and Robin Hendry, "Philosophy of Chemistry", *The Stanford Encyclopedia of Philosophy (Winter 2011 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2011/entries/chemistry/>, accessed November 4, 2012, see subsection 6.

[115] VCH Publisher's Encyclopedia of Physics begins their entry on chemical bonding by citing the union of physical and chemical theory. It was written by a chemist: Henry A. Bent, "Chemical Bonding", *Encyclopedia of Physics, Second Edition,* editors R. G. Lerner and G.L. Trigg (New York: VCH Publishers, 1991), 146. Also, physicist John Gribbin refers to chemistry as a branch of physics. See: John Gibbin, *Q is for Quantum...,* 79. In addition, physicist Leonard Susskind holds that chemistry is a branch of physics. See: Leonard Susskind, *The Cosmic Landscape*, (New York: Back Bay Books, 2006) 8-9, 173. In contrast, Thompson cites philosopher John Dupre, philosopher Nancy Cartwright and physiologist Alan Garfinkel when asserting that chemistry and physics do not reduce. Thompson, 440.

cites philosophers John Dupre and Nancy Cartwright, as well as physiologist Alan Garfinkel.[116]

I will not attempt to adjudicate this dispute here. However, I will assert that: in order for Thompson to be successful in pointing to the regenerative capacities of cells (i.e. being multiply realizable) as evidence against epiphenomenalism, the dispute as to whether chemistry reduces to physics would have to be settled definitively in favor of the philosophers who deny inter-theoretic reduction. Given the trend of science to treat chemistry as a branch of physics (*and the many associated successes in applied science*), victory for the dissenting philosophers seems 'nowhere in sight' if we define victory as buy-in from the scientific community at large accompanied by compelling alternate explanations for the physical-chemical reductive behaviors discovered thus far.

Even if one was inclined to accept the scientifically unpopular view that chemistry does not reduce to physics, Thompson's use of cellular multiple realizability does not dispel the more pressing worry of epiphenomenalism with respect to human minds. If all of my cells were to mysteriously stop regenerating at some instant in time, I would, no doubt, eventually die. Nonetheless, there would be some short period during which I would be an autonomous, conscious agent. The regenerative aspects of my body at any given instant are not what enable my consciousness at that moment. It is, therefore, hard to see how my own multiple realizability entails that my consciousness is not epiphenomenal.

[116] Thompson, 440.

Perhaps Thompson has something more subtle in mind in the case of brains and human cognition. When he asserts the macro-level organization of complex systems is not epiphenomenal due to multiple realizability, perhaps – in the context of human minds – he does not mean the regenerative replacement of cells in brain and body. From his text there is no way to tell since he does not consider the specific example of human minds in support of his multiple realizability assertion. He limits his consideration to cells. In regard to human minds, we could give a charitable interpretation of his intent by considering the Thompson quote I presented above in regard to co-emergence associated with brains. We might say that the whole in the form of global dynamic patterns of large-scale integration of neuronal activity is multiply realizable in the sense that it is continually regenerated by new versions of its constituent local neuronal activities. In this way we avoid the problem of consciousness continuing for some short time even after the regenerative processes at the level of cells has stopped. By taking this more charitable interpretation, it is obvious that the large-scale integration of neuronal activity (i.e. the cognitive whole) would cease immediately if the local neuronal activities stopped – regardless of whether the cells engaging in those activities continue self-regeneration. In this way, the existence of the whole is dependent on its continual regeneration via local neuronal activities, and this regeneration is a kind of multiple realizability across time.

Does the sort of multiple realizability I am describing above somehow entail that the global structure it supports is not epiphenomenal? I do not see how. We are still left with the same question we had in the case of the cell, namely: Is the large-scale global behavior reducible without remainder to the local behavior? If the global neuronal activity is reducible without remainder, then the global structure is epiphenomenal. If it is not, and the irreducible remainder can exert influence on the local neuronal activities, then downward causation is something immanent within the processes of brains and (presumably) their associated minds.

The brain is indeed a complex system, whereby the global structure of large-scale integrated neuronal activity is nonlinearly related to small changes in localized neural activities. Those localized behaviors multiply through branching causal chains of neural activities, most of which are processed by neural interconnect structures shaped by life-long chains of experiences. We would like to know whether the global behavior is entirely reducible to the behavior of the local constituents. If it is reducible, there is no place for purported new causal powers of emergent global structures. However, even if we are unable to reduce it, Thompson and his strong emergence advocates are not out of the woods. The inability to reduce could well stem from the extreme nonlinearity of the system. In other words, the irreducibility could be an epistemic limitation as opposed to an ontologic feature of brains and minds. This is where Thompson's argument for strong emergence with respect to human mind loses steam. Presently, there is no way of knowing for certain whether the purported irreducibility of the human mind is epistemic or ontologic. Though patterns of large-scale integration in neural activity might, in some sense, be multiply realizable because they are continually regenerated through local neural activities, we cannot claim this multiple realizability entails the global structure is not epiphenomenal without *assuming* complex dynamic systems are inherently irreducible without remainder. Assuming a remainder amounts to assuming some new causal power has emerged. This is of course precisely what Thompson wants to show. If we can only get there by assuming the answer, the question has been begged.

3.2.c. Differing True Counterfactuals and Holistic Claims

Thompson offers a second argument in favor of the non-epiphenomenal status of global dynamic structure. He bases it on an observation that true counterfactuals associated with macro-level systems can be different than true counterfactuals associated with their micro-level processes. I can communicate his idea no more succinctly than he does:

> "The organization [of the macrostructure] is necessary for certain subsequent events, but the particular constituents are not. If the organization had been different and those constituents the same, the events would not have occurred; but if the constituents had been different (within certain limits) and the organization the same, the events would still have occurred (because of the way the constituents were organized).* Given that different counterfactual statements are true of the microlevel constituents and the macrolevel organization, the two cannot be identified."[117] [Bracket content added for clarity. Parenthetical content in original. *A footnote reference was not quoted to avoid confusion with footnotes in this paper.]

Thompson expresses this argument in explicit response to Kim's view that talk about a system's organization having emergent causal influence is really just a macro description of microlevel causal events.[118] In Kim's view, there are no new causal powers that come into being as a result of a system's organization – at least none that are not already accounted for by the causal properties and processes of the micro

---

[117] Thompson, 436-437.

[118] Thompson, 436.

constituents. Thompson views Kim's claim as a claim of epiphenomenalism with respect to a system's organization.

There is a problem with Thompson's argument in that the differing true counterfactuals with respect to wholes and their constituents are no guarantee against the organization of a system being epiphenomenal – if 'epiphenomenal' if is taken to mean 'no new causal powers'. For example, Thompson's reasoning can be applied as well to my car as to complex dynamic systems. The organizational structure of my car is necessary for me to travel from my home to the university. If the pistons were in the passenger's seat rather than the engine block, my arrival at the university would not have occurred. However, if the individual parts of the car had been different (within the engineer's +/- tolerances) and the organizational structure the same (i.e. pistons where they belong) I would have arrived at the university. In fact, the parts of my car are regularly changed out when I take it to the dealership for service. Therefore, we could use Thompson's argument to support a claim of strong emergence with respect to my car – i.e. my car has new causal powers that are not explained by the causal properties and processes of its parts. This is, of course, not true. The engineers who designed my car used their knowledge of the causal properties and processes associated with its parts to determine the causal powers of my fully assembled vehicle. Nowhere in their design equations did they enter a term to account for the onset of new causal powers once the assembly was complete.

A Thompson defender might argue that my car is not a complex dynamic system nonlinearly related to its constituent parts or processes, and my car does not exhibit circular causality of operational closure to norms of sustenance. Yet this counter misses the point. Thompson asserts true counterfactuals associated with global processes of the whole are different from true counterfactuals associated with processes of the constituent parts. He concludes these differences demonstrate the whole cannot be causally identified with the parts. Yet my car is evidence that his conclusion is incorrect. Does the conclusion somehow become correct when applied to complex dynamic systems? I think not. For it to be correct, one would need to show that the claimed irreducibility of complex dynamic wholes is not a matter of mere epistemic limitation associated with the sheer complexity and nonlinearity of the complex system. One would have to make an assumption amounting to a leap of faith that such apparent irreducibility is an ontologic feature of the system as opposed to an epistemic limitation of our finite ability to examine it. Such a leap of faith not only assumes the result it purports to demonstrate but also runs headlong into conflict with our knowledge of nature and physics. I will explain why shortly.

3.2.d. Nonseparable Hamiltonian Physics Argument

As Thompson advances his argument against Kim, he complains that Kim's consideration of emergence is limited to examples of early British emergentism, and these examples predate the contemporary science that ought to guide our thinking.[119] This is a fair criticism in the sense that Kim does not analyze dynamic systems complexity. Nonetheless, it loses force when we look more carefully at Thompson's view of modern science and his hopes that science negates the overdetermination worry.

---

[119] Thompson, 438.

Thompson's modern science view relies on a Hamiltonian formulation of physics rather than a Newtonian formulation. As he notes, the Hamiltonian formulation holds energies are fundamental in nature rather than forces.[120] There is little wonder as to why he would be attracted to this formulation of physics. Under it, the total energy of a physical system becomes nonseparable when the equations of motion describing the system include nonlinear terms. In other words, the total energy of the system cannot be expressed as the sum of the energies of the system's constituent parts.[121] To Thompson and other similar-minded philosophers, the nonseparability of the Hamiltonian (i.e. nonseparability of the total energy) signals a kind of holism in dynamic complex systems. These holistic systems cannot be reduced to their parts. Yet, they are deterministic.

Thompson cautions that his view of 'deterministic' should not be taken to mean the future is fixed by the present status of the world. He notes that real-world empirical experience with nonlinear dynamic systems will always display some amount of stochasticness – i.e. random behavior often referred to as 'noise'. When coupled with the sensitivity of nonlinear systems to tiny changes in initial conditions, that noise causes the future behavior of the system to be unpredictable. Because of that unpredictability, Thompson believes we should view our application of determinism to nonlinear dynamic systems as a mere analysis technique and not as an ontologic statement about what exists in nature.[122] Under this view, determinism is more or less a convenient way to

---

[120] See Thompson's Appendix B language referencing footnote 8 on page 438 as well as footnote 8 on page 480.

[121] Thompson, 429.

[122] Thompson, 430-431.

approximate what will happen in the very near future; however, it is not a real feature of the world.

This is a curious qualification regarding determinism. I find it curious for two reasons. First, as Thompson prosecutes his claim for Newtonian-based determinism being a mere analysis technique, he asserts determinism is only useful in describing the behavior of systems free of noise. The assertion implies that, without the random noise behavior, the systems *would* behave completely deterministically. On what grounds then would we deny ontologic status to determinism without also denying ontologic status to stochasticism? Are they both mere analytic techniques that have nothing to do with what is real in nature? While it is true they are both techniques to predict or describe what we see in nature, the behavior they delineate is *real*. There is a very real sense in which – at any moment in a dynamic system – what happens next is determined by the current status of the system's components in terms of momentum, chemical composition, the surrounding milieu of fields and forces, the laws of nature *and* any momentary energy contribution due to stochastic noise. That the more distant future behavior of the system as a whole is not entirely predictable due to the noise in nature and the nonlinearity of dynamic complexity in no way casts doubt on the ontological status of determinism. Put another way, determinism is not somehow depleted of its ontological status merely because we have discovered some natural systems (i.e. dynamically complex systems) that are so sensitive to noise as to be unpredictable in the long run.

The second reason why I find Thompson's declaration about determinism curious has to do with his unequal treatment of the Hamiltonian. In the case of Hamiltonian analysis he is quite prepared to accept that the analytical nonseparability of the system's total energy signals a genuine ontological feature of the natural world – i.e. a holism with respect to dynamic systems.[123] He gives no reason for accepting the reality of the nonseparability of the Hamiltonian while denying the same to determinism. Yet, describing physical systems using the Hamiltonian function is clearly an analytic technique.

Furthermore, Thompson is applying this analytic technique to slow, large-scale objects such as cells, immune systems, and brains – as opposed to relativistic or quantum scale objects such as atoms, electrons and photons. In other words, the context of analysis is entirely classical physics as opposed to relativistic and/or quantum physics. This is an important distinction because the Hamiltonian equations of motion are *equivalent* to Newton's equations of motion when applied to nonrelativistic, non-quantum systems.[124] Given this equivalence, we could use Newtonian mechanics to analyze the systems. Under Newton's equations, the energy of the system *is* separable into the energies of its constituent parts. The integral of the motion of a classical system (i.e. its total kinetic energy) is the sum of the kinetic and potential energies of its constituents. Similarly, the total mass of the reagents participating in a chemical reaction (i.e. its total

---

[123] Where Thompson describes the Hamiltonian origin of holism in dynamic systems on page 430, he seems to hedge his support of it by describing it as a proposition of Robert Bishop without explicitly stating he agrees. However, later on page 438 he claims that modern science should guide our thinking about emergence and, in a related footnote 8 on page 480, Thompson makes clear that his notion of modern science includes the nonseparability of the Hamiltonian. Additionally, he presents the concept of holism and the Hamiltonian on pages 428 and 429 in the context of defending against the very notion of Kim's downward causation.

[124] R. H. Good Jr., "Hamiltonian Function", *Encyclopedia of Physics (Second Edition)*, Rita G. Lerner and George L. Trigg (ed.), (New York: VCH Publishers, 1991), 476-477.

potential energy) equates to the sum of the masses of the products of the reaction.[125] Given these facts, Thompson should conclude that nonseparability of the Hamiltonian is the result of an analysis technique and *not* an ontologic feature of nature.

Of course, the Newtonian equations of motion become impossibly impractical to wield when applied to the individual constituents of nonlinear complex systems because complex systems have millions of moving parts. This impracticality is, at least in part, why other analysis techniques have been employed to describe dynamically complex systems. Nonetheless, *any* analysis technique that is able to more simply approximate the behavior of a nonlinear complex system will inevitably be unable to reduce the system's behavior at any instant in time to the distant-past behavior of the system's components. This inability is due to the contribution of stochastic noise amplified by the nonlinearity of the system. Given this fact, there are no grounds for declaring the nonseparability of the system's total energy under Hamiltonian analysis signals the ontological status of holism. In other words, even though the system displays global behavior that we are unable to reduce to the behavior of constituent parts, the irreducibility is an epistemic limitation born of stochasticism and non-linearity. It does not necessarily signal the emergence of holism with new causal powers born of complexity. The holism Thompson promotes is, at best, a conjecture, and more likely a mirage born of the analysis technique. Thus we are still left with Kim's overdetermination worry in that the behavior of the dynamic whole is, in principle, explainable by the behavior of its constituent parts were it not for nature's interfering noise amplified by the system's complex nonlinearity.

---

[125] Lincoln Wolfenstein, "Conservation Laws", *Encyclopedia of Physics, Second Edition,* editors R. G. Lerner and G.L. Trigg (New York: VCH Publishers, 1991), 177.

Given this chain of causal explanation, there is nothing left for the purported new causal powers of holism to contribute to the behavior of the system's constituents without over-determining that behavior.

3.2.e. Disavowing Nature's Reducibility in Favor of New Causal Powers

Thompson's ill-founded ascription of holism to nonlinear complex systems leads him to conclude the macro-level characteristics of a system can be micro-dependent and micro-involving without being analytically reducible into the entities and properties of the system's micro-level constituents.[126] Were it true, holism would support his claim that the macro-level structure of these systems is not epiphenomenal, and the micro-level interactions happen the way they do because of the emergent causal powers of the macro-level characteristics.[127] This stance places him in direct conflict with the founding principles of Kim's overdetermination argument. These are (in paraphrase):

1. Every emergent must be determined or dependent on a physical event or property (i.e. its emergence base).
2. The causal powers of an emergent are identical with or a subset of the causal powers of its emergence base (i.e. no new causal powers).
3. All causes of physical events are themselves physical (i.e. causal closure).[128]

Thompson rejects 1 and 2 because they are "an expression of physicalistic ontological reductionism."[129] Under his alternate view, stochastic noise prevents macro

[126] Thompson, 438.

[127] Thompson, 436.

[128] Thompson, 435.

[129] Thompson, 439.

behavior from reducing to micro constituents, and therefore holism in the form of new causal powers must be in play. Yet, as I have pointed out, under a broader view, global behavior might well be determined by the combination of both the noise and the properties of the basal components – making global behavior ontologically reducible though not epistemically reducible in the case of nonlinear dynamic complexity.

Thompson also disavows physicalistic ontological reductionism on the basis of a non-hierarchical world view. His notion of co-emergence in complex dynamic systems does not comport with emergents being necessarily determined by lower level events. They are, instead, reflexively co-determined (sometimes at the same level), whereby whole emerges from constituent parts while parts emerge from whole. He believes a particular view of modern science – affirming holism in nonlinear complex systems – demonstrates the causal powers of the whole are not identical to or a subset of the constituent parts. They, instead, include *new* causal powers not present in the parts. In fact, he disparages the notion of upward determination from basal conditions:

> "This image of nature as a mereologically ordered hierarchy grounded on a base level of particulars is a metaphysical picture projected onto science, whereas the image science projects is of networks of processes at various spatiotemporal scales, with no base-level particulars that "upwardly" determine everything else."[130]

Though Thompson cites other philosophers in support of this view, the view suffers from a kind of parochialism in that it draws conclusions using an excessive focus on

[130] Thompson, 439.

dynamic complexity at the human timescale. It overlooks the bigger picture of nature as a thirteen billion year progression. While true that science has uncovered networks of processes at various spatiotemporal levels, the image that science projects for all of nature is one of upward evolution from lower-level constituents. The evidence for this lies in one of science's greatest successes, namely: the evolutionary view of the universe from a big bang some 13.7 billion years ago.[131] Thompson's '…networks of processes at various spatiotemporal scales…' were not always here. Therefore, to hold that such a network of processes is the image of reality projected by science is parochialism of the here-and-now kind. Unless we are to ignore the vast body of evidence pointing to the big bang, we can hardly embrace the notion that there was no upward determination from base-level particulars – the particulars at the big bang. These are the very particulars from which Thompson's networks of processes have evolved. We can argue as to whether the fundamentals are forces or energies (i.e. better described by Newtonian versus Hamiltonian analysis); however, the resolution of that argument (either way) will do nothing to support an image of networks of processes with no historic determination upward from fundamental particulars.

The majority of serious scientists do not dispute that: Human consciousness came after immune systems. Immune systems came after cells. Cells came after molecules. Molecules came after atoms. Atoms came after protons. Protons came after photons and electrons.[132] Photons and electrons came after the "...primordial plasma of nature's

[131] Stephen Hawking and Leonard Mlodinow, *A Briefer History of Time,* (New York: Bantam Dell, 2005) 68.

[132] Hawking, 68-72.

elementary constituents."[133] The view that all of nature has evolved from fundamental particulars at the earliest moments of the big bang is more than a mere armchair theoretic projection by scientists. It has provided testable predictions, and it has led to empirically verified discoveries at both extremes of the spatiotemporal scale – from prediction and discovery of the universe-wide microwave background residue left by the big bang[134] to prediction and discovery of various subatomic particles via particle accelerators.[135] Given the coherence of the big bang view and its many empirical successes, Kim's first two principles cannot be dismissed with shaky conjectures of holism and a parochial view of nature as disjointed processes.

What about the third principle founding Kim's overdetermination argument – i.e. the causal closure of the universe? Thompson claims to have abided by that principle in his theories of part/whole reflexive co-emergence – *provided that causal closure is not narrowly construed as causal closure of the microphysical*.[136] The circular causality that Thompson describes in, for example, his promotion of Varela's cellular autopoiesis holds that the causal inter-dependence of a biological cell's cognitive self and its cognitive environment is operationally closed and causally independent of the microphysical milieu in which it is situated. It may be thermodynamically dependent on the microphysical milieu for energy; however, that energy is not, in Thompson's view, the cause of the emergent cellular structure and processes. The emergent operational closure of the cell and its cognitive environment (i.e. the exterior influences that are meaningful to its

---

[133] Greene, *The Fabric ...*, 247.

[134] Greene, *The Elegant Universe*, 348-350, and Hawking, 73.

[135] Greene, *The Fabric... , 252-256.*

[136] Thompson, 439-440.

sustenance) is a relational process instead of an up/down dependency. Thompson views nature as an irreducibly, relational processes at all levels.[137]

> "In the process view, "up" and "down" are context-relative terms used to describe phenomena of various scale and complexity. There is no base level of elementary entities to serve as the ultimate "emergence base" on which to ground everything."[138]

In addition to the objections that I have already raised against this view on the grounds of unsupported conjectural holism and big bang evidence against nature-as-disjointed-processes, I will make one final observation. The driving force behind Thompson's arguments is a noticing that certain types of systems (namely, nonlinear dynamic complex systems) display macro-level behaviors that are not readily explained by observing the behavior of their constituents. They are seemingly irreducible. We are therefore tempted to suppose that something about the dynamic structure of the whole brings into being new causal powers that did not exist prior to the whole coming into being. There is a narrow sense in which this is true. My car has the power to cause me to arrive at the university; whereas its individual parts sitting in bins at the factory do not. However, this is *not* the sense in which Thompson and others are ascribing new causal powers to dynamic complexity. The 'new' causal powers of my fully-assembled car are fully explained by and reducible to the combined causal powers of its individual components. In contrast, under the Thompsonian view of dynamic complexity, the 'new' causal powers are not so explained or reduced. They are *in addition* to the causal powers

---

[137] Thompson, 440.

[138] Thompson, 441.

of the individual components. There are two ways to interpret this claim. One is disputed by science, and the other is a question for science instead of philosophy.

The first interpretation of 'new causal powers' would be one in which the energy associated with those powers comes into being as a result of the emergent structure of the whole. This interpretation proposes the impossible because it conflicts with the principle of conservation of energy. Since Isaac Newton's day we have known that energy is neither created nor destroyed. Every physical interaction requires an even trade of energy. Any apparent loss or gain of energy is ultimately a conversion of one form of energy to another. Conservation of energy has withstood every paradigm shift in physical theory since Newton – including relativity and quantum mechanics. As of yet, no experiment has contradicted energy conservation.[139] Its abandonment would lead to a wide variety of inconsistencies in natural theory, none of which are supported by genuine inconsistencies in nature. Therefore, 'new causal powers' cannot mean new energy sources coming into being.

The other interpretation of 'new causal powers' would be a sort of nomological reordering under dynamic complexity, whereby new energy is *not* created. Instead, the new causal powers rely on existing energy converted from another form due to some sort of shift in the law-like regularity of cause and effect when dynamic complexity is present. We could debate how or why such a shift would occur or whether it is even possible; however, in one sense the debate is mere mental gymnastics – or at least premature effort.

---

[139] Greene, *The Fabric...*, 426.

Why? Because the existence of new causal powers as nomological reordering is measurable since the onset of these powers requires an uptake/conversion in energy.

Consider, for example, the phenomenon of Rayleigh-Bénard convection cells. If we were to gradually heat a still pool of oil from the bottom, the narrow time interval during which the fluid first breaks into a pattern of hexagonal cells is the time when new causal powers are said to come into being. In order for those new causal powers to, in fact, *have* power to do anything that is not already accounted for by the activity of the underlying constituent parts, there must be an accompanying uptake and/or conversion of energy – unless we succumb to the superstition that energy can be created. When the new causal powers come into being, the fluid must take up an incremental amount of energy from its surrounding environment. That amount of energy must be equivalent to the genuine work done by the new causal powers and unaccounted for by the existing causal powers of the underlying constituents.

In the case of Rayleigh-Bénard cells, the available energy sources are somewhat limited. The energy must come from the chemical bonds of the fluid or from the thermal energy of the surrounding environment since the phenomenon we are considering is not occurring under conditions conducive to nuclear reaction, nor is it occurring in a system in motion relative to a large body nearby – i.e. sourced via gravitational energy.[140] Therefore, we should either witness a change in the chemical composition of the fluid displaying Rayleigh-Bénard convection, or we should witness an increased cooling at one

[140] Power is work per time and we have available in nature only four possible forces to accomplish that work: gravitational, weak nuclear, strong nuclear and electromagnetic. See Hawking, 120. Chemical energy and the mechanical energy of thermodynamic collisions or vibrations both rely on electromagnetic forces binding atoms and molecules. Having eliminated the uptake of energy in the form of nuclear or gravitational, only chemical and/or thermodynamic remain as candidates for our Rayleigh-Bénard cells.

or more of the surfaces of the fluid as it uptakes thermal energy from the surrounding environment during the onset of its purported new causal powers.

To my knowledge, no one has attempted to look for the *onset* of new causal powers associated with dynamic complexity by conducting chemical and/or thermodynamic measurements. Perhaps the energy changes associated with these new causal powers would be so minuscule as to be immeasurable with current technology. On the other hand, the daunting task of measuring minuscule bits of energy transfer has not dampened the enthusiasm for experiments aimed at detecting the *retreat* of causal powers by attempting to detect a change in mass and/or thermal energy when organically complex systems like humans stop functioning. These experiments have been conducted by a variety of – mostly unqualified – researchers bent on witnessing the purported exit of the soul upon death.[141] Under the law of conservation of energy, if souls were to exist in a way that allows them to exert causal powers over the body, they too would have to obey nature's laws. Their exit from the body would cause a loss of energy measurable as loss of mass or heat or some other energy form. When experiments aimed at detecting this energy loss were conducted by soul-advocates and the results showed no such loss, the experiments were often discredited by the pseudo researchers themselves as being flawed or inconclusive. Nonetheless, any hint of energy perturbation was hailed as genuine evidence of the soul's departure – sometimes with comical embellishment.[142]

---

[141] Mary Roach, *Spook: Science Tackles the Afterlife*, (New York: W. W. Norton & Company, 2005) 79-106.

[142] See Roach, 93-94 for reasoning about calculated soul volume signaling leprechauns are likely discarnate humans.

To be clear (and charitable), I am not suggesting that belief in new causal powers associated with complex dynamics is superstitious or on par with soul searching. While soul searching is the sort of topic that discourages legitimate peer reviewed research facilities from consuming precious resources, emergence is not. The topic of new causal powers is respectable enough to garner legitimate resources *if* advocates are prepared to press for the necessary experimentation. Therefore, we can and should expect advocates of emergent causal powers to exhaust available scientific methods in justification of their position. My point is this: since these new causal powers must adhere to the same laws as all other natural phenomena, our philosophical debate about their nature and efficacy is, in some sense, premature. We could make better use of our labor by looking to science for confirmation of their existence before we wrangle over their role in emergence – less we discover we are debating miracles.

CHAPTER 4

WHAT DOESN'T WORK, WHAT DOES, AND HOW WE MIGHT USE IT TO MAKE PROGRESS

4.1. Talking Past Kim

Prior to this point, I have been defending Kim's view that strong emergence characterized by claims of new causal powers leads to a legitimate concern about overdetermination. My arguments have countered Thompson's claim that science supports the holism associated with new causal powers. They have also defended causal closure of the universe and the evolution of nature from fundamental particulars. Yet, my criticism of Thompson's holism and his non-fundamentalist view of nature should not be mistaken for a dismissal of all aspects of his stance on autopoiesis or the enactive approach to embodied mind. Some of Thompson's proposal does work with respect to easing the polemics surrounding the origins of cognition, and that portion of his proposal can be used to make progress on the problem of mental causation. Thompson eases the debate on origins of cognition by focusing our attention on meaning-giving as signaling the incipient stage of cognition. His approach has potential to relieve at least some of our concerns over epiphenomenalism in that it hints of a way to parse out our often-undisciplined lumping of several concepts under the broad conceptual heading of 'consciousness' – e.g. mind, cognition, wakefulness, awareness, sub-consciousness, intentionality, etc. We will see how shortly.

My disagreement with Thompson engages where he interprets autopoiesis and the enactive approach as signaling holism and a non-fundamentalist physics – and where he attempts to use these controversial concepts to dismiss Kim's worry over downward causation and overdetermination. Thompson's holism and non-fundamentalist physics are questionable for the reasons I have provided above. Yet even if my arguments against holism and non-fundamentalism were completely incorrect, his application of those concepts would not succeed in overcoming the threat of epiphenomenalism at the level of mental *awareness*. With or without holism and non-fundamentalism, Thompson's argued conclusion does not work in that it merely shifts the debate from talk of mind/body to talk of information/causation. As we shall see, this shift *does* succeed in overcoming part of Kim's worry over downward causation; however, it only does so by banishing the threat of epiphenomenalism at the non-conscious, cellular level of mind and not at the level of mental awareness. Furthermore, the progress it makes in relieving a portion of the mental causation worry is not due to (or reliant on) holism or non-fundamental physics. I will explain the reasoning behind these assertions in detail. First, let's consider how Thompson shifts the debate.

Thompson's notion of reflexive co-emergence supposes the parts emerge from the whole as much as the whole emerges from the parts – this in lieu of unidirectional emergence, whereby wholes come from lower-level preexisting parts. In the case of human nervous systems, he envisions the parts as local neuronal processes (instead of the neurons themselves), and he envisions the whole as the global structure of neuronal processes.[143] Thus rather than a mind/body or software/hardware Cartesian view, mind is the entire operationally-closed system (whole and parts) arising out of the physical milieu of the nervous system and its interactions with the world. Up versus down with respect to part-whole versus whole-part become context relative.[144] On this view, we could think of part and whole as operating at the same level while together arising out of (and drawing energy from) the microphysical milieu. The threats of downward causation and overdetermination are seemingly bypassed since the microphysical milieu is merely the energy source for the operationally-closed, co-emergent system. Nonetheless, we should ask whether Thompson has addressed Kim's worry or replaced it with another.

In essence, Thompson talks past Kim's issue. Kim wants to know how the mind can exercise causal efficacy on the physical in ways that are not already explained by activities of the physical constituents underlying and correlating to mental activity. Thompson dismisses this sort of talk as symptomatic of an inflexible Cartesian world

---

[143] Thompson, 423.

[144] Thompson, 441.

view.[145] Yet, even if Thompson is correct regarding his own alternative explanation involving reflexive co-emergence, it is difficult to see how that alternative answers Kim. It seems, instead, to talk about something different. Furthermore, nowhere in his own view does Thompson explain how the nonphysical is able to exercise causal efficacy over the physical – though he stealthily relies on it to do so. How so? He replaces the mental-to-physical causation mystery with an information-to-physical causation mystery. Let's reconsider his argument focusing this time on his use of the concept of information.

Perhaps Thompson is correct in that – when we worry about minds having genuine causal efficacy on the physical – we are hopelessly stuck in a Cartesian trap. Maybe we do need to view mind as embodied, and perhaps we should allow that emergence involves operational closure through circular causality. Nonetheless, this alternative, non-Cartesian view suffers from a gap no less daunting than the more traditional mind/body gap. That gap has to do with information and causation, and it becomes apparent in Thompson's discussion of the enactive approach.

During the early stages of his book, Thompson explicates the enactive approach to embodied mind by making a distinction between 'informational stimulus' and 'physical stimulus'.[146] The idea is that, in complex dynamic systems, information stimulus serves as a triggering condition rather than an efficient cause. In other words, rather than being the primary reason for change or producing/generating effects, information stimulus merely triggers a shift in the global structural activities of the complex system. It is a necessary element even though it is not sufficient to account for the global behavior.

[145] Thompson, 438.

[146] Thompson, 69-70.

Thompson further explains that information counts as information relative to the organism. The organism as a whole makes the stimulus what it is (i.e. information) by reacting only to stimulus having some positive or negative bearing on the organism's sustenance.

If we follow the convention of thinking of mind as the upward direction and its effects as the downward direction, Thompson's view can be seen as a claim that the thing caused in the downward direction is information. Yet this does not answer Kim's particular worry. The fact that the information is caused is precisely why the subsequent physical action – responding to the information – occurs. Whether we think of information as a triggering condition or, instead, an efficient cause of action has little bearing on solving the mental causation issue. The physical response of the organism to the stimulus does not occur without the stimulus having the status of information. Kim's mind/body downward causation problem has been replaced by an information/action causation problem.

Consider an example. If a careless hunter points a loaded gun at a fellow hunter, avoidance behavior is triggered in the other hunter. If, on the other hand, the careless hunter points a loaded gun at a human toddler, no avoidance behavior is triggered. Whether we say the adult's avoidance behavior was due to 'mind' or 'information stimulus' makes little headway in resolving the mental causation issue. Thompson has simply rephrased the problem. Instead of, "How does [mind] cause action?" we are left to ask, "How does [stimulus becoming information] cause action?" Under Thompson's view, we see our notion of mind is replaced with physical stimulus becoming information. Regardless of whether we view the associated causal efficacy as same level or downward, a question still remains as to the physical causal efficacy of a nonphysical entity. In one case the non-physical entity is mind. In the other it is information.

4.2. Dual Realms Approach to Mental Causation

Nonetheless, Thompson has made progress on the problem of mental causation. He argues that the incipient stage of cognition occurs at the level of biological cells where operational closure of the organism gives meaning to stimulus relevant to sustenance. In doing so, Thompson focuses our attention on an aspect of cognition which is clearly not epiphenomenal. Consider again the careless hunter example. In light of the differing reactions that a toddler and an adult would give to a loaded gun, the meaning-giving associated with an adult's reaction is far from being merely epiphenomenal.

On the other hand, some thinkers might object that, in the careless hunter example, the meaning giving is a matter of physical neural conditioning via lived experience. In other words, the mental awareness of danger was brought about by physical sensations causing physical changes in neural networks such that the awareness is not necessary for the avoidance behavior. Take, for example, trained soldiers in combat. They can be so conditioned to avoid line of fire that they do so without even being aware of it. Therefore, instances exist in which conscious awareness is not needed for the causal outcome to obtain.   This objection points to an area of cognition where conscious awareness and meaning-giving come apart while both still remain within the conceptual boundary of mind.

Thompson's contribution to the debate focuses on and follows meaning-giving down to its most primitive level, far below any possible threshold where awareness could be present. At the level of biological cells, the physical conditioning of experience plays no role in the inborn meaning-giving that sustains a cell at the moment of its origin and throughout its life. We do not need to accept Thompson's over-reaching notions of holism or non-fundamental nature to find value in his core strategy of placing the incipient stages of cognition at the meaning-giving level of biological cells. We can employ his notion of cognition in cells to help parse out the confusing array of concepts that are often associated with mind. Incipient cognition as raw meaning-giving isolates the follow-on concepts of intentionality, consciousness and awareness from the list of attributes necessary for cognition. Cells do not perceive or have intentions, nor are they conscious in the sense of possessing wakefulness, nor are they aware. Yet they react to certain stimuli in ways that seem mindful of maintaining their operational integrity.

Microbiologist Michael Denton performs a splendid job of illuminating just how spectacular the many and varied self-regulatory activities of the cell are when he performs a thought experiment magnifying a living cell to the size of a city.[147] At that scale, atoms would be the size of tennis balls and observation of the automated processes of self-maintenance would lead one to conclude that cells are more complex and better organized than the most advanced human factories. Furthermore, the exquisite choreography of the atoms and molecules within the cell while they flawlessly restore energy reserves, eliminate waste and protect from invasion invite us to conclude that

[147] Michael Denton, "The Puzzle of Perfection", *Readings in the Philosophy of Religion [Second Edition]*, editor Kelly James Clark (Peterborough, Ontario, Canada: Broadview Press, 2008) 74-82.

mindfulness of some sort must be involved. If we were to witness these behaviors in a modern robotic factory, we would conclude that mind is involved at least in the design. Denton uses this thought experiment to conclude that the mindfulness involved is beyond the reach of evolutionary chance and, instead, signals intelligent design. We need not – and I think should not – go so far in speculating about its origins. Nonetheless, Denton's thought experiment can serve as a compelling argument for Thompson's claim of the incipient stage of cognition occurring in biologic cells – i.e. giving meaning to certain stimuli by selectively and appropriately responding only to those that have significance to organism's norms of sustenance.

If we accept Thompson's assertion, we may conclude that cognition – in its most primitive form – need not involve the awareness we humans enjoy. This is a liberating realization because it allows us to parse our answer to the question of mental causation according to our parsing of progressively higher levels of mind. The meaning-giving associated with cognition at the level of cells is *not* epiphenomenal. Its causal efficacy is evidenced, for example, by the survival of the cell. On the other hand, it is not so clear that meaning-giving associated with awareness possesses powers of efficient cause. At that higher level, the advocates of epiphenomenalism have some justification for claiming meaning-giving is a matter of physical neural conditioning via lived experience. Beneath each conscious mental choice there is a neural network doing the choosing, and that network was formed and grew via a lifetime of experience – beginning with the formation of our concepts during infancy and even pre-birth.[148]

[148] Empirical evidence suggests that prenatal human fetuses learn something about hand and face via touch while still in the womb. M. Myowa-Yamakoshi and H. Takeshita, "Do Human Fetuses Anticipate Self-

Put another way, the remarkable ability of cells to choose from among stimuli in ways appropriate to their ongoing maintenance does *not* involve consciousness or awareness. Yet it *does* involve incipient cognition in the form of meaning-giving. This meaning-giving has genuine causal efficacy not explained by underlying layers of the physical world. In contrast, our human awareness is, at least in part, explained by underlying layers of the physical world – i.e. neural correlates. If so, awareness exercises no causal powers (*of the productive/generative kind*) that are not already accounted for by the underlying physical/neural correlates. The real causal work occurs at the pre-aware level of those correlates. This is *not* a claim that awareness has no effect on the status of the world. I will explain this distinction shortly.

Consider again my Figure 4. 'The Experience of Mental Cause' as presented in Chapter 2 (pg.32). Figure 4 suggests our folk sense of mental causation might just be the experience of the transitional, 'in-between' physical-neural activities involved when one neural state causes another. Under this view, the *experience* of mental causation does not display downward causation. Keep in mind that Figure 4 was presented prior to considering Thompson's work – i.e. before we allowed the possibility that the incipient stage of cognition could occur at the level of cells. If we allow this possibility, an enhanced version of Figure 4 presents an interesting way to think about mental causation. Here is a schematic diagram.

Oriented Actions? A Study by Four-Dimensional (4D) Ultrasonography" *Infancy* 10(3) (2006), 289-301. Additional empirical evidence suggests that prenatal humans learn something about flavors based on their mother's diet during pregnancy. J. A. Mennella, C. P. Jagnow, and G. K. Beauchamp, "Prenatal and Postnatal Flavor Learning by Human Infants" *Pediatrics* Vol. 107 No. (6 June 2001).

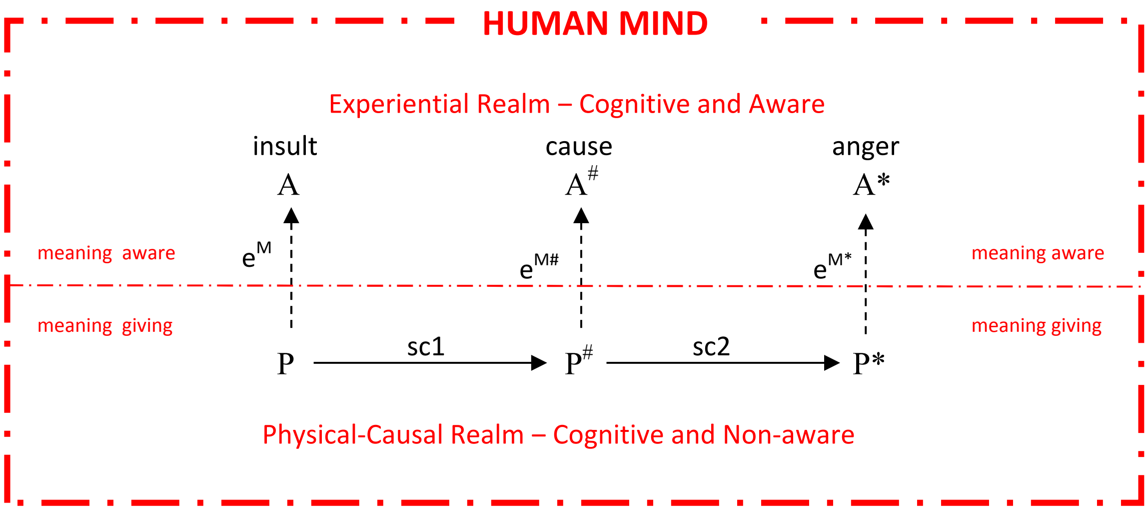

Figure 6. Dual Realms of Meaning Giving and Meaning Aware

Neurons are living cells and thus perform Thompsonian meaning-giving in the sense of responding appropriately to stimuli that positively or negatively affect their sustenance. These stimuli include the signaling they receive from other neurons. From the lower portion of Figure 6, the meaning-giving associated with a given neuron's operational closure occupies the physical-causal realm of mind. The stimuli that neurons choose from and the meaning-giving they participate in is vastly distributed across a thousand or more connections to different neurons – i.e. synapses.[149] Via these connections the neuron can operate in both an excitatory and an inhibitory role with respect to sending and receiving 'information' in the form of electrical potentials to and from other neurons.[150] The connections are not random accidents. This fact is often conveyed by the informal neuroscience maxim, "Cells that fire together wire together."[151]

[149] Researchers estimate the adult human brain has roughly one hundred billion neurons, each sharing (on average) one thousand connections to other neurons. J. M. Schwartz and S. Begley, *The Mind and The Brain* (New York: Harper Collins, 2002), 105, 111.

[150] Schwartz and Begley, 106.

[151] Schwartz and Begley, 107.

In other words the strength and spatial extent of a neuron is determined by which information stimuli it selects from in terms of responding with excitement or, instead, responding with inhibition to other neurons. Since a neuron's response to other neurons influences not only its own strength but also its own spatial extent, we can consider this response to be Thompsonian meaning-giving at the incipient stage of cognition. It is a first order meaning-giving on which the higher order meaning-giving of perception and mental representation depend. This dependence could be a reason why genuine cognition via artificial intelligence is so elusive.

Put another way, a neuron's responses to other neurons determine the manner in which it grows and is sustained. Therefore, those responses to stimuli from other neurons qualify as Thompson's sustenance-involving meaning giving. The maxim 'cells that fire together wire together' takes on the dimension 'cells that fire together cognize together'. Since a given neuron's meaning-giving response extends outward to hundreds or thousands of other neurons, it has a direct effect on its connection community. The meaning-giving is causal because its connection community is affected by its signaling in the same way that it is affected by that community. Furthermore, a neuron's connection community can ultimately serve as the efficient cause of muscle movements controlling actions. Therefore, the pre-conscious group cognition of interconnected neurons is causal. Thompson has given safe harbor to a limited notion of mental causation, and that causation is not downward.

This safe harbor is limited to cellular meaning-giving and does not directly extend to conscious awareness – though it has an indirect link in that it functions as the underlying basis of the higher order meaning giving associated with fully aware reasoning. Here is where we might be departing from at least one portion of Thompson's enactive approach to embodied mind. In particular, Thompson's fifth core principle of the enactive approach asserts: experience is central to explaining mind and not merely an epiphenomenal side issue.[152] Certainly any full explanation of mind must account for the role of experience. On the other hand, the claim that experience is not epiphenomenal is not certain *depending on what Thompson means by 'experience'*. If, in this context, we take 'experience' to mean mere interaction with the world, we have no conflict with Thompson. An organism's interactions with the world are not epiphenomenal in that those interactions have a deterministic effect on its follow on behavior, its sustenance, and the world. If, on the other hand, we take 'experience' to reference only awareness, his fifth core principle is less certain. Though not epiphenomenal in the sense of making no difference, awareness is not *necessarily* involved in the mental causation associated with cognition. To see why, consider, for example, how we decide on and perform a wide variety of actions without engaging awareness. These might include:

- scratching our nose in response to an itch, or
- shifting in our seat to reestablish comfort, or
- choosing our words to match our audience, or
- slowing our vehicle in response to a curve, etc.

[152] Thompson, 13.

Of course, many times these sorts of actions are performed with full engagement of our awareness and full attention to the sensations associated with them. Yet awareness and attention need not and often are not so engaged.

Blind sight patients provide even starker evidence against awareness being required for mental causation. These brain-injured people have limited or no awareness of their own capacity to see the world. They do not experience vision. Yet they can perform many tasks in ways that can only be accomplished if visual sensations are present and cognitive processing of those sensations is occurring.[153] This is clear evidence that awareness is not essential to cognition or mental causation.

Consider yet another studied example of mental causation without awareness. In 1999 neurophysiologist, Benjamin Libet, published results from an ingenious neuroscience experiment.[154] In that experiment, Libet measured tiny electrical changes on the scalps of human subjects near the skull's vertex in an area known to produce 'readiness potential' signals. Readiness potentials occur approximately 550 milliseconds (ms) *in advance* of subjects performing simple willed acts such as quick flexion of the hand or wrist. These readiness potentials are a reliable predictor of impending movement and therefore signal a subject's decision to move. Libet asked test subjects to perform a simple hand flexion at any time of their choosing without any outside stimulation or prompting. The subjects were also asked to note the precise time at which they *became aware* of their decision to perform the movement. This was done by having them

---

[153] Tim Bayne, *The Unity of Consciousness*, (Oxford: Oxford University Press, 2010) 98-99.

[154] Benjamin Libet, "Do We Have Free Will?", *Journal of Consciousness Studies*, Vol. 6, August/September, Imprint Academic, Thorverton U.K., 1999, 47-49.

continually observe a fast-moving ‘clock’ indicator. The clock was nothing more than a dot orbiting a numbered dial every 2.56 seconds. For each test, the Readiness Potential was recorded and displayed in relation to the clock time and also in relation to the physical onset of the subject’s hand movement. The subject’s reported time of awareness regarding their decision was also plotted on the timeline.

As could be expected, the subject’s verbally reported moment of decision-awareness occurred before the onset of hand movement. In fact, it occurred on-average about 200 ms before the actual movement. However, that moment of awareness was also 350 ms or more *after* the Readiness Potential. The sequence was: 1) readiness potential, 2) awareness of decision, and 3) movement of the hand. In other words, the brain sends the signal for the hand to move well before the subject is aware they have decided to move. Furthermore, the results were repeatable across test subjects. Libet’s experiments are further confirmation that mental causation can occur without awareness.

Nonetheless, Figure 6 is not, by any stretch, a complete depiction of mental causation, meaning giving and awareness. It ignores the categorical differences associated with the progressively higher orders of meaning giving beginning with the differential response to stimuli at the cellular level and becoming progressively more sophisticated in perception, mental representation, and finally reasoning. In this way Figure 6 reflects a weakness in Thompson’s view. Thompson offers little explanation as to how the meaning giving that he sees in biologic cells evolves into the sort of meaning giving that philosophers employ in their writings or artists employ in their paintings. If Thompson’s meaning giving at the level of cells is true, then it is likely to be only a first order phenomenon that serves as the basis for distinct progressively higher order levels of meaning giving.

In regard to mental causation, Figure 6 is incomplete because it portrays mental awareness as radically epiphenomenal – i.e. having no effect whatsoever on the physical world. Though I have argued via examples that awareness is sometimes not necessary for cognition and decision making, I am not claiming it has no effect on the world. Epiphenomenalism is usually employed to describe some feature of the world that: a) has no causal powers and b) produces no physical effect on the world.[155] Mental awareness *does* have an effect on the world; however, it does so without being the productive/generative cause of that effect. As I indicated earlier, my use of 'cause' with respect to mental causation is meant as efficient cause in the productive/generative sense. I am not using 'cause' in the passive, non-generative sense of true counterfactuals. For example, if I happen to look in a mirror while in a sunlit room, and I discover ink on my nose, I could use 'cause' in the passive true counterfactual sense of: "The mirror caused my discovery." In other words, had the mirror not been there, I would not have made the discovery. This passive sense of cause is not the sort of thing that scientists are worried about when they claim that something non-physical cannot generate motion or other energy changes in something physical. On the other hand, if I mean to employ 'cause' in the productive/generative sense of efficient cause when I consider the ink spot on my nose, I cannot claim the mirror caused my discovery. Instead, the solar nuclear reactions that produced the streaming sunlight were the productive/generative phenomenon that caused me to discover the ink. The presence of the mirror was merely a necessary

[155] Blackburn, 117.

circumstance for the sunlit to do the causal work of my discovering the ink. Through this distinction, we can see that some things can affect the world in the true counterfactual sense while not being the productive/generative cause of that effect. This distinction is important to interpreting the dual realms approach that I communicate via Figure 6.

In Figure 6, I have relabeled Figure 4's higher level 'M's as 'A's to signify that the higher level is *awareness* and not the more general concept of mental. Instead, the full contents of the box labeled 'MIND' are mental – including the meaning-giving activities of cells in response to stimuli. Note that any or all of the three 'A's in Figure 6 could – under some circumstances – be missing without changing the outcome or entailing that cognition is missing. For example, you may have had the experience of leaving a conversational encounter and feeling slightly angry without knowing why – only to later realize that you were insulted by the subtly condescending speech of your interlocutor. At the time of the original encounter you were unaware of both the insult and the cause of the anger. Yet the meaning-giving necessary to be insulted and to cause anger was there and fully engaged at some level below awareness. Similarly, you may find yourself engaged in a conversation about politics with a friend and be genuinely surprised to hear her exclaim, "Well, you don't have to get angry about it!" You were unaware of your own irritation until it was pointed out to you. In these examples, the nonessential nature of the awareness signals its lack of causal efficacy, while the mental causation riding beneath is performing the work of efficient cause. In other words, we need not assert that, for example, the *awareness* of insult is downwardly causing the anger since the insult is occurring and causing the anger at some cognitive level beneath awareness.

An antagonist to my proposal might object that I handpicked examples to fit the needs of my argument and that these sorts of cognition-without-awareness episodes are not the norm. I did handpick the examples; however, I'm not so certain they do not reflect the norm in that we cannot be aware of the cognition we are unaware of. When we reflect on our experiences at the end of each day, that mental summary hardly seems to account for 16 hours of cognition. It is an open question as to how much of the day's cognition was 'forgotten' and how much occurred without our awareness – and even this distinction is itself unclear.

An antagonist might further point out that very often we are completely aware of mental transitions in ourselves and others and we act on those transitions by endorsing them or criticizing them. These actions have a clear causal outcome on the status of the world. Therefore, my Figure 6 representation is open to strong objection since I portray mental awareness as having no causal efficacy.[156]

This objection is appropriate in that Figure 6 does not account for the influence that awareness has on the outcome of our interactions with the world. On the other hand, I am stopping short of proposing awareness is epiphenomenal in the radical sense of having no effect whatsoever. Instead I am proposing (via Figure 6) that awareness has no causal powers in the productive/generative sense of efficient cause. I am not proposing that it has no effect on the outcome of our cognitive activities in the true counterfactual sense. When we reason, it seems obvious that our awareness of our thoughts has an effect on the conclusions we reach. It seems that our awareness of a proposition reflects back into our reasoning so as to affect the conclusion. This is somewhat analogous to a

[156] I am indebted to Bernard Kobes for this objection.

mirror reflecting light. However, this reflective role of awareness might be the result of underlying mirror-like neural feedback as expressed by the axiom ‘cells that fire together wire together’. In other words, even with respect to awareness, the causal work in the productive generative sense is still occurring at the level of the physical correlates and the experience of awareness is causal only in the passive, true counterfactual sense of: “If I hadn’t been aware of the proposition’s contents I wouldn’t have reasoned to the correct conclusion.” Under this view, the causal closure of the physical world that our empirical sciences demand is still preserved – provided that we stick to the productive/generative sense of efficient cause.

My point is this: by combining Thompson's view of incipient cognition via meaning-giving at the cellular level with the notion that the *experience* of mental causation is something entirely different than the mental causation itself, we can parse mind into meaning giving and meaning aware realms. Under this dual realms approach, mental causation survives as efficient cause at the lower level without invoking the need for downward efficient causation from the nonphysical realm of mental awareness to the concrete world of neural physics and chemistry. Our intuitive sense of awareness being causally involved in choice is dampened but not destroyed by the realization that its causal role is true counterfactual as opposed to productive/generative.

One could object to the dual realms approach by focusing on my acceptance of Thompson’s cornerstone claim that the incipient stage of cognition extends down to the cellular level. The objection might go as follows. Thompson’s claim of cognition in cells points to a group of behaviors each of which help to sustain the organism such that it can function as a system within some band of conditions and over some duration of time. These individual behaviors include, for example, formation of a protective wall, restoring energy reserves, elimination of waste, combating invasion, etc. Yet, when considered in isolation, each behavior can be explained by genetically encoded instructions originating as accidents of nature preserved via natural selection. When viewed by itself, a particular behavior need not be seen as involving meaning-giving. The happenstance of several of these behaviors being encoded in a single, self-replicating species does not entail meaning-giving. In order to assert otherwise, one must provide an explanation as to exactly how many self-sustaining behaviors are needed to achieve a threshold whereby meaning-giving (and therefore cognition) is in play.

This argument should sound familiar. Recall that, when we explored B-weak emergence in Chapter 1, I raised a concern over its fuzzy boundary. I asked how many different incarnations of glider guns qualify that phenomenon as multiply realizable and therefore B-weak emergent. Analogously, how many different incarnations of self-sustaining behaviors qualify a system as giving meaning to its environment and therefore cognitively emergent? As was the case with glider guns, the answer is unclear. Nonetheless, a fuzzy boundary need not disqualify either phenomenon from multiple-realizability. Just as a sorites puzzle with respect to grains of sand does not challenge the ontological status of heaps, a sorites puzzle with respect to meaning-giving in the form of self-sustaining behaviors does not challenge the ontological status of cognition. One who challenges Thompson on the notion of cognition in cells only postpones the inevitable. At some level of complex self-sustaining behavior the challenger must admit to the acceptability of a fuzzy categorical border for the onset of meaning-giving or they must deny the existence of meaning-giving altogether. Accepting a border at the threshold of life allows us to speak coherently about the existence of mental causation without requiring that nonphysical properties like awareness possess causal powers of the productive/generative kind in violation of nature's physics.

By considering the question of how many self-sustaining behaviors qualify an organic dynamically complex system as giving meaning (and therefore possessing cognition) we realize that Mark Bedau may have been too hasty in dismissing his own notion of B-weak emergence from the mind/body debate.[157] If we accept my suggestion to embrace Thompson's view of cognition emerging as meaning-giving at the cellular level while denying his holism and non-fundamental physics, then the attributes of cognition line up with the hallmarks of B-weak emergence. The behaviors of cells are both ontologically and causally reducible to their micro facts. Nonetheless, the ways in which those behaviors can be self-sustaining is multiply realizable. Just as Conway's game-of-life glider guns are realizable in an indeterminate number of configurations, the self-sustaining behaviors for complex dynamic systems are realizable in an indeterminate number of configurations.

Along these same lines, recall our earlier exploration of B-weak emergence, whereby the indeterminacy associated with the number of ways an emergent can exist signals computational irreducibility. In other words, there is no algorithm for predicting how many different ways a particular emergent phenomenon can occur. Though the instances of the phenomenon are reducible to its basal conditions, the full set of its different

[157] Bedau, 161.

instantiations is not deducible from those bases. No method exists for determining how many different ways the B-weak emergent phenomenon may occur – except for simulation by aggregation of the basal components or some facsimile and letting the system run.

In the case of game-of-life glider guns, computational irreducibility is due, in part, to the infinity of the cellular plane and therefore the infinity of possible initial configurations. We could never try them all to find whether a new, undiscovered form of glider gun exists. Analogously, the computational irreducibility of complex dynamic systems with respect to the emergence of different possible self-sustaining behaviors is due, at least in part, to the sensitivity to initial conditions that nonlinear systems possess in concert with the influence of stochastic noise. The randomness of noise in terms of amplitude and pulse width applied to the sensitivity to initial conditions prevents us from predicting the system's global behavior into the future. This, in turn, prevents us from discovering every possible way in which self-sustaining behaviors might occur. Additionally, we have no shortcut such as an algorithm for determining the total count. Therefore cognition as meaning-giving in the form of self-sustaining behaviors is computationally irreducible.

In sum, cognition as meaning-giving supporting closed loop self-sustenance has all the attributes of Bedau's B-weak emergence. Its boundary is fuzzy; however, its meaning-giving behaviors are multiply realizable. For a given instance of self-sustaining behavior we can trace a causal reduction path back down to initial, lower-level conditions of atoms and molecules. On the other hand, there is no corresponding deduction path that leads us upward from the basal conditions to discovery of the full range of possible meaning giving behaviors. We have realized this status of B-weak emergence by resisting Thompson's notions of holism and non-fundamental physics while accepting his explanation of incipient cognition at the cellular level. To repeat Bedau's own assessment: "This kind of robust weak emergence reveals something about reality, not just about how we describe or explain it. So the autonomy of this robust weak emergence is ontological, not merely epistemological."[158]

Some thinkers might resist my demotion of the emergence involved in cognition from strong emergence to B-weak emergence and my demotion of awareness from efficient causal to true counterfactual causal. They might claim that my stubborn resistance to holism and my insistence on reducibility is radical. They might point out that attempts to radically reduce nature to a handful of physical particulars fail to account for the richness of our experience of the world. Yet my position is not one of radical reductionism. I recognize that, even the most radical reductionist, must find something brute at the end of the reduction path. My acceptance of B-weak emergence amounts to acceptance of many phenomena that are brute and unaccounted for solely by the physical particulars at the moment of the big bang. As the big bang particulars evolved according to efficient causation, they uncovered ontologically autonomous phenomena such as the class glider gun in cellular automata and meaning-giving in the wide and persistent class of self-sustaining behaviors in nature. These emergent phenomena are not reducible as a class despite the reducibility of their instances. Acceptance of B-weak emergence is acknowledgement that radical reductionism cannot explain all that we see. It is a

[158] Bedau, 183.

reasoned way to coherently accept that there are metaphysically brute phenomena without abandoning what science tells us about the conservation of energy and the closure of the physical world with respect to productive/generative causes.

Consider one final note about the fuzzy boundary where self-sustaining behavior qualifies as meaning-giving. This boundary occupies the very intersection of our two senses of information – i.e. the colloquial sense of meaning versus the Shannon technical sense of amount. On the one hand we are asking, "When does the processing of stimuli amount to the organism giving meaning to the ocean of data imparted by its milieu?" Here we are asking about information in its parochial sense of meaning. On the other hand we are asking, "How many different kinds of self-sustaining behaviors must an organism display before we can claim it is giving meaning to the ocean of data imparted by its milieu?" Here we employ information in its technical sense of 'amount' to answer a question about the onset of its parochial sense of 'meaning'. As with the comparative case of glider guns and hopper guns, the Shannon number – as a measure of information – can be used as a yardstick to measure our level of relative certainty with respect to assertions about cognition in the form of meaning-giving. In other words, counting self-sustaining behaviors as a measure of cognition might be a useful analysis tool in our exploration of the boundary conditions for questions about, for example:

- when a dynamic complex system is alive,
- when (if ever) a computer has genuine cognition,
- when a fetus is a human, and
- when a brain-injured human is no longer a person.

CHAPTER 5

SUMMARY CONCLUSION

The dual realms approach to mental causation offers the prospect of retaining a concept of mental causation without the worry of downward efficient causation from nonphysical conscious awareness to physical neural constituents. The path to this approach has been somewhat arduous. We began in Chapter 1 by exploring Jaegwon Kim's bipartite, weak-versus-strong emergence distinction and followed by considering an important middle ground offered by Mark Bedau's tripartite view. Without this middle ground, Kim too hastily relegates all claims of emergence into one of two bins, namely: incoherent metaphysics or unremarkable platitudes about our epistemic limitations. I have argued that Kim’s claim of incoherence does not stand against close examination because it improperly attempts to force-fit logical entailment onto interpretations of strong emergence claims. Yet there are at least two phenomena purported to instantiate strong emergence that do not rely on logical entailment – one involving inductive inference and the other abductive reasoning to the best explanation (§ 1.2.). I have also argued that a more careful consideration of Bedau's computationally irreducible B-weak emergence is both coherent and genuinely worthy of remark, for it illuminates an objective ontology of macro-level structures with irreducible overarching identities (§ 1.3. and § 1.4.).

In Chapters 2 and 3, we gave more careful consideration to Kim's worry over downward causation and associated overdetermination. Kim notices that, in a causally closed universe, the purported new causal powers characterizing strong emergence have no function – at least none that is not already performed or explained by the causal powers of the basal constituents. Thus any claim of downward efficient causation from emergents to bases (e.g. mind to body) leads to overdetermination (§ 2.1.).

After examining the nature of Kim's downward causation worry, we considered Thompson's enactive approach to explaining embodied mind. That approach views mind as emerging from dynamic complexity and circular causality, whereby mind and cognitive environment reflexively co-emerge. In Thompson's view, co-emergence signals a kind of holism, whereby the global structure of the emergent whole brings with it new causal powers that are not present in or reducible to the constituent parts (§ 2.3. and § 2.4.). From that vantage point, Thompson defends the notion of *synchronic* downward causation using quantum non-locality and *diachronic* downward causation using Hamiltonian analysis, whereby the system's energy is not separable into the energies of its constituents. I have argued that his defense fails due to three faults: 1) It makes an improper assumption about the role of causality when quantum nonlocal systems decohere (§ 3.1.a.-d.), 2) It ignores Einstein's relativistic ban on instantaneous transmission of information across physical distances (§ 3.1.e.-f.), and 3) It fails to recognize that – at nonrelativistic speeds and non-quantum physical dimensions – Hamiltonian analysis is equivalent to Newtonian analysis, thereby implying holism is not an ontologic feature of the world but, instead, the result of an analysis technique (§ 3.2.d.). These arguments relegate synchronic downward causation to the status of incoherence. Causes must come before their own effects.

In regard to *diachronic* downward causation, Thompson mounted additional defense by claiming that new causal powers associated with strong emergence must, indeed, exist due to the multiple realizability of, for example, living organisms. The fact that all parts of a living thing are replaced over its lifetime implies the existence of global causal powers that hold the process of the particular living-self invariant. Thompson argued these global causal powers must be new in that they do not reduce to the causal powers of the organism's constituents because chemistry does not reduce to physics. In response, I have argued that claims of irreducibility between chemistry and physics are far from accepted by mainstream science. Furthermore, they conflict with the long trend of empirical evidence pointing to the reducibility of chemistry to physics (§ 3.2.b.). Additionally, I have argued against the bald assumption that dynamic complexity is ontologically irreducible (as opposed to merely epistemically irreducible), because this assumption amounts to question begging with respect to claims of new causal powers – i.e. the assumption assumes the conclusion being sought.

As another defense of diachronic downward causation, Thompson tried to show new causal powers driving downward causation are evidenced by the fact that true counterfactuals associated with complex wholes are different than true counterfactuals associated with their constituents. In response, I have argued this counterfactual incongruence is not evidence of new causal powers because the same counterfactual incongruence is true of my car. Yet, my car's global structure and process behavior is entirely reducible to its parts and their properties (§ 3.2.c.).

In his final assault on Kim's downward causation worry, Thompson claims modern science portrays the world as a network of different processes operating at different spatial-temporal scales – rather than a hierarchic framework reducible to fundamental particulars. Therefore, 'up' and 'down' are relative terms, and reducibility is not a necessary feature of nature. I have argued that such an implausible view of nature is in direct conflict with most of what big bang cosmology has taught us about nature. Therefore, the price of justifying downward causation by accepting such an improbable view of nature is simply too high (§ 3.2.e).

In Chapter 4, I noted how Thompson's treatment of Kim's downward causation worry talks *past* Kim and shifts the problem from mind/body causation to information/physical causation (§ 4.1.). I then suggested we might make progress by accepting the more palatable aspects of Thompson's view regarding the incipient stage of cognition, and combining those aspects with my earlier observation about the distinction between mental causation and the *experience* of mental causation. Following that suggestion, I developed this combination into a dual realms approach to mental causation, whereby mind is modeled as having both a meaning-giving and a meaning-aware realm (§ 4.2.). This approach allows for a more believable notion of embodied mental causation without requiring downward causation in the productive/generative sense from the nonphysical realm of awareness to the physical realm of neurons and muscles. Nonetheless, I argued that the dual realms approach need not disqualify awareness from having a causal impact on the status of the world provided that its causal role is viewed as passive true counterfactual. Finally, I located the notion of incipient cognition at the cellular level squarely within the conceptual realm of Bedau's B-weak emergence –

provided we resist Thompson's holism and non-fundamental physics. Mental causation of the productive/generative kind is alive and well at the level of biologic cells while mental awareness avoids radical epiphenomenalism provided we temper our claims about its role in causation to that of true counterfactual.

REFERENCES:

Aristotle. *Aristotle Introductory Readings*, translated by Terence Irwin and Gail Fine. Indianapolis: Hackett 1996.

Bayne, Tim. *The Unity of Consciousness*. Oxford: Oxford University Press, 2010.

Bedau, Mark A. “Downward Causation and Autonomy in Weak Emergence”. *Emergence, Contemporary Readings in Philosophy and Science,* editors Mark A. Bedau and Paul Humphreys. Cambridge: The MIT Press, 2008, pp.155-188.

Bennett, Karen. “Why the Exclusion Problem Seems Intractable and How, Just Maybe, to Tract It”, *Nous*, 37:3 (2003): 471-497.

Bent, Henry A. "Chemical Bonding". *Encyclopedia of Physics, Second Edition,* Editors R. G. Lerner and G.L. Trigg. New York: VCH Publishers, 1991.

Blackburn, Simon. *Oxford Dictionary of Philosophy, second edition.* Oxford: Oxford University Press, 2005.

Callahan, Paul. “What is the Game of Life?”. Math.com. http://www.math.com/students/wonders/life/life.html. (Accessed 12/16/12.)

Cornell University. “Conway’s Game of Life”. Cornell’s Math Explorer’s Club. http://www.math.cornell.edu/~lipa/mec/lesson6.html. (Accessed 12/16/12.)

Denton, Michael. "The Puzzle of Perfection". *Readings in the Philosophy of Religion [Second Edition]*, Editor Kelly James Clark. Peterborough, Ontario, Canada: Broadview Press, 2008, pp.74-82.

Falcon, Andrea. "Aristotle on Causality". *The Stanford Encyclopedia of Philosophy (Fall 2011 Edition)*, Edward N. Zalta (ed.). http://plato.stanford.edu/archives/fall2011/entries/aristotle-causality/. (Accessed September 29, 2010.)

Fodor, Jerry. “Special Sciences (Or: The Disunity of Science as a Working Hypothesis)”. *Emergence, Contemporary Readings in Philosophy and Science,* editors Mark A. Bedau and Paul Humphreys. Cambridge: The MIT Press, 2008, pp.395-409.

Good, R. H. Jr. "Hamiltonian Function". *Encyclopedia of Physics (Second Edition)*, editors Rita G. Lerner and George L. Trigg. New York: VCH Publishers, 1991.

Greene, Brian. *The Elegant Universe.* New York: W. W. Norton & Company, 1999.

Greene, Brian. *The Fabric of the Cosmos.* New York: Random House, 2004.

Gribbin, John. *Q is for Quantum, An Encyclopedia of Particle Physics.* New York: The Free Press, Simon & Schuster, 1998.

Gribbin, John. *Schrödinger's Kittens and the Search for Reality.* Boston: Little, Brown and Company, 1995.

Hawking, Stephen and Leonard Mlodinow. *A Briefer History of Time.* New York: Bantam Dell, 2005.

Horgan, Terry. "Materialism, Minimal Emergentism, and the Hard Problem of Consciousness". *The Waning of Materialism,* editors G. Bealer and R. Koons. Oxford: Oxford University Press, 2010.

Hume, Dave. *An Enquiry Concerning Human Understanding* 1748. *Central Readings in the History of Modern Philosophy, Descartes to Kant,* Second Edition, editors Robert Cummins and David Owens. Belmont, CA: Wadsworth Publishing Company, 1999, p. 351.

Jackson, Frank. "Epiphenomenal Qualia". *The Philosophical Quarterly*, Vol. 32, No. 127 (Apr., 1982), pp. 127-136.

Kim, Jaegwon. *Essays in the Metaphysics of Mind.* Oxford: Oxford University Press, 2010.

Laplace, Pierre Simon marquis de. *A Philosophical Essay on Probabilities,* translated by Truscott and Emory. New York: John Wiley & Sons, 1902.

Libet, Benjamin. "Do We Have Free Will?". *Journal of Consciousness Studies*, Vol. 6, August/September, Imprint Academic, Thorverton U.K., (1999): 47-49.

LifeWiki. "Welcome to LifeWiki, the wiki for Conway's Game of Life." http://conwaylife.com/wiki/Main_Page. (Accessed 3/3/13.)

Martin, Edwin. "Conway's Game of Life".
http://www.bitstorm.org/gameoflife/standalone/ (Accessed 3/3/13.)

McGraw-Hill Dictionary of Scientific and Technical Terms, Sixth Edition. New York: McGaw-Hill, 2003.

Mennella, J.A., C. P. Jagnow, and G. K. Beauchamp. "Prenatal and Postnatal Flavor Learning by Human Infants". *Pediatrics,* Vol. 107 No., (6 June 2001).

Myowa-Yamakoshi, M. and H. Takeshita. "Do Human Fetuses Anticipate Self-Oriented Actions? A Study by Four-Dimensional (4D) Ultrasonography". *Infancy,* 10(3) (2006): 289-301.

Nadeau, R. and M. Kafatos. *The Non-Local Universe*. Oxford: Oxford University Press, 1999.

Roach, Mary. *Spook, Science Tackles the Afterlife*. New York: W. W. Norton & Company, 2005.

Rosenlum, B. and F. Kuttner. *Quantum Enigma.* Oxford: Oxford University Press, 2006.

Schwartz, J.M. and S. Begley. *The Mind and The Brain.* New York: Harper Collins, 2002.

Susskind, Leonard. *The Cosmic Landscape, String Theory and the Illusion of Intelligent Design.* New York: Back Bay Books, 2006.

Thompson, Evan. *Mind in Life - Biology, Phenomenology, and the Science of Mind*. Cambridge: The Belknap Press of Harvard University Press, 2010 paperback, 2007 original.

Von Baeyer, Hans Christian. *Information, The New Language of Science*. Cambridge: Harvard University Press, 2003.

Weisberg, Michael, Paul Needham, and Robin Hendry. "Philosophy of Chemistry". *The Stanford Encyclopedia of Philosophy (Winter 2011 Edition)*, Edward N. Zalta (ed.). http://plato.stanford.edu/archives/win2011/entries/chemistry/. (Accessed November, 4, 2012.)

Wolfenstein, Lincoln. "Conservation Laws". *Encyclopedia of Physics, Second Edition,* Editors R. G. Lerner and G.L. Trigg. New York: VCH Publishers, 1991.