

The Perspectival Nature of Emergence

Elanor Taylor

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Philosophy.

Chapel Hill
2012

Approved By:

Marc Lange

John Roberts

William Lycan

Laurie Paul

Karen Neander

Abstract

ELANOR TAYLOR: The Perspectival Nature of Emergence (Under the Direction of Marc Lange)

Philosophers use the term “emergence” for the phenomenon whereby a whole has properties that are in some way autonomous from, while still dependent upon, the properties of its components. Emergence is crucial for those who hold that there are distinct levels of properties in nature, because it is thought to be the relation that accounts for distinctions between such levels. However, there is a problem that afflicts philosophical accounts of emergence: for any purported case of emergence, there is a micro-level property that apparently makes the emergence “collapse.” Given this micro-level property, the purportedly emergent macro-level property is no longer autonomous, and there seem to be no good grounds for excluding the collapse-inducing property from the set of micro-level properties on which the macro-level is dependent.

I examine various attempts to solve this problem of “collapsing emergence” by imposing principled restrictions on the admissible micro-level properties underlying cases of emergence. I argue that such attempts fail and that to solve the problem we must appeal to the connection between emergence and explanation. I argue that we use the concept of emergence to track the availability (or not) of scientific explanations, and that what qualifies as an available explanation differs across different “perspectives”, where a “perspective” is the point of view of a particular observer. This connection between emergence and

explanation provides a solution to the collapse problem because, I argue, collapse-inducing properties fail to provide resources for the explanations that emergence tracks, and so cannot render a macro-level property non-emergent. Embracing this solution (and hence the possibility of emergence) involves embracing the idea that emergence is relative to a perspective.

Understanding emergence as relative to a perspective is contrary to some popular conceptions of emergence, which take emergence to be a metaphysical, perspective-independent phenomenon. I argue that such conceptions are misguided and that emergence *can* play a major role in philosophical and scientific enquiry into relationships between properties of parts and wholes and between different levels in nature, but it can do so *only* as a perspectival phenomenon, rather than a metaphysical one.

For John A. Taylor

Acknowledgements

Thanks are due first of all to my dissertation committee. Marc Lange was a model advisor. Through his example and guidance I have realized just how hard it is to do philosophy and also how worthwhile. Insightful feedback and criticism from John Roberts and William Lycan proved invaluable in shaping my views. I am grateful to Laurie Paul for an introduction to the practice of metaphysics, and to Karen Neander for her feedback and support in the later stages of this project.

UNC has a thriving philosophical community and it has been a wonderful environment in which to pursue a PhD. Thanks are due particularly to my graduate student colleagues Clair Morrissey, Kristen Bell, Emily Given, Emily Kelahan, Katrina Elliott, Cathay Liu and Jamin Asay for their friendship and philosophical companionship. Nadine Langford and Kerry McKenzie provided long-standing friendship across the Atlantic, and without Patrick Connolly's patience and strength this project would never have been completed.

Throughout my time in this program my mother held seemingly unshakeable faith in my abilities, and her curiosity and intellectual vigor are a life-long inspiration. My stepparents, Ron and Val, provided love and support as did my brother Dawson and my stepbrothers and stepsisters. Finally, I am forever grateful to my father for encouraging my interest in philosophy at an early age. This dissertation is dedicated to his memory.

Table of Contents

Introduction	pg 1
Chapter 1: The Idea of Emergence	pg 4
1. Introduction	pg 4
2. Dependence and Autonomy.....	pg 4
3. The Historical Development of the Concept of Emergence.....	pg 10
4. Emergence in Contemporary Philosophy and Science	pg 15
5. Associated Ideas.....	pg 18
6. Problems for Accounts of Emergence	pg 25
7. Moving On.....	pg 30
Chapter 2: Collapsing Emergence.....	pg 31
1. Introduction	pg 31
2. The Collapse Problem	pg 32
3. Distinctions to Hitch To	pg 45
Chapter 3: An Appeal to Explanation	pg 56
1. Introduction	pg 56
2. Emergence and Explanation	pg 56
3. A Minimal Condition for Explanation	pg 67
4. A Solution to the Collapse Problem	pg 74

Chapter 4: An Explication of Emergence.....	pg 84
1. Introduction	pg 84
2. Philosophical Approaches to the Interpretation Question.....	pg 85
3. The Concept of Emergence in Scientific Practice	pg 90
4. The Need for a New Approach.....	pg 96
5. An Explication of Emergence.....	pg 100
6. Considering Carnap's Criteria	pg 104
7. Objections	pg 113
8. Putting the Explication to Work.....	pg 115
9. Conclusion.....	pg 123
Chapter 5: The Perspectival Account of Emergence.....	pg 125
1. Introduction.....	pg 125
2. The Positive View	pg 125
3. Metaphysical Implications	pg 130
4. Differentiation Among Emergents	pg 138
5. Downward Causation	pg 142
6. Objections and Replies	pg 145
7. Conclusion	pg 150
Summary and Conclusion	pg 151
Bibliography	pg 157

Introduction

“The ultimate coherence of the notion of an emergent property remains controversial...”¹

“The higher quality emerges from the lower level of existence and has its roots therein, but it emerges therefrom, and it does not belong to that lower level, but constitutes its possessor a new order of existent with special laws of behavior. The existence of emergent qualities thus described is something to be noted, as some would say, under the compulsion of brute empirical fact, or, as I should prefer to say in less harsh terms, to be accepted with the “natural piety” of the investigator. It admits no explanation.”²

Philosophers use the term “emergence” for the phenomenon whereby a whole has properties that are in some way autonomous from, while still dependent upon, the properties of its components. The position that some macro-level properties are emergent is fairly popular: emergence has been presented as the relationship that obtains between different levels in nature, such as the relationship between properties of qualitative experience and physical properties³, and some argue that a successful account of emergence can vindicate the apparent causal autonomy of certain higher-level entities, particularly the entities of the special sciences⁴. Emergence is crucial for those who hold that there are distinct levels or

¹ Kim, J. (1995) pg 224

² Alexander, S. (1920) pg 47

³Chalmers, D. (2006); Chalmers, D. (1996)

⁴ Wilson, J. (forthcoming)

layers of properties in nature, because emergence is thought to be the relation that accounts for distinctions between such levels.

I am going to argue that this conception of emergence is misguided. Existing accounts of emergence face a problem, *the collapse problem*, which shows that emergence does not track relations between distinct levels of properties in nature, as it is a metaphysically arbitrary relation. However, I will argue, this does not mean that emergence is *completely* arbitrary. Instead, according to the solution I will offer to the collapse problem, emergence tracks the availability or otherwise of scientific explanations.

I will develop a positive account of emergence, *the perspectival account of emergence*, according to which a property is emergent if and only if a certain scientific explanation is unavailable to an observer. I will argue that, in addition to avoiding the collapse problem, the perspectival account of emergence makes for a unified and naturalistic account of emergence.

In Chapter 1 I will offer an introduction to the idea of emergence, a survey of its historical development and an examination of some related concepts. In Chapter 2 I will present the *collapse problem* for accounts of emergence, and will discuss some failed responses to that problem. In Chapter 3 I will offer a solution to the collapse problem, based on the connection between emergence and scientific explanation. In Chapter 4 I will develop and defend an explication of the concept of emergence as the unavailability of a scientific explanation for an observer, based on the findings of Chapter 3. And finally in Chapter 5 I will formulate and defend the *perspectival account of emergence*, based on the explication developed in Chapter 4. I hope to show that, in light of the perspectival account of

emergence, the “*ultimate coherence of the notion of an emergent property*” is no longer controversial.

Chapter 1: *The Idea of Emergence*

Section 1: Introduction

In this chapter, I will offer a brief introduction to emergence in Section 2, an overview of the historical development of the concept of emergence in Section 3, and an overview of contemporary approaches to emergence in Section 4. In Section 5 I will examine some closely related concepts, including reduction, physicalism and supervenience, and in Section 6 I will discuss objections to accounts of emergence, including the problem of downward causation and the claim that a commitment to emergence is non-naturalistic and non-parsimonious. This discussion will set us up for Chapter 2, in which I will present the *collapse problem* for metaphysical accounts of emergence.

Section 2: Dependence and Autonomy

The term “emergence” is used for a diverse range of phenomena – so diverse, in fact, that it can be hard to recognize what such phenomena have in common. For an illustration of this variety, consider the following list, every item on which has been described by a philosopher or a scientist as a case of emergence:

1. The relationship between convection rolls in a body of heated fluid and the micro-physical states of that fluid⁵
2. The relationship between configurations in Artificial Life worlds and the cells that form them⁶
3. The relationship between phenomenal and neural properties⁷
4. The relationship between entangled states and the components of their constituent systems⁸
5. The relationship between the properties of a chemical compound and the properties of its constituent elements⁹

Despite the variety of the phenomena in question, however, the philosophical literature on emergence isn't entirely heterogeneous. Discussion tends to gather around three central questions, which are:

- Interpretation: What is emergence?
- Viability: Is there any emergence?
- Taxonomy: What is the relationship among the different forms of emergentism?¹⁰

⁵ Kelso, J.A. Scott (1995)

⁶ Bedau, M. (2003); Dennett, D. (1991)

⁷ Chalmers, D. (1996)

⁸ Redhead, M. (1995); Teller, P. (1986)

⁹ Broad, C.D. (1925)

¹⁰ Where "emergentism" is the position that there are emergents.

In Chapters 2-5, I will develop answers to all three of these questions, but before doing so I will offer a brief overview of some central features of the concept of emergence.

Despite the diverse range of different accounts and purported cases of emergence, there is a common, schematic conception of emergence: assuming a distinction between micro-level and macro-level properties, emergent properties are macro-level properties that are in some sense both dependent on and autonomous from their underlying micro-level properties.¹¹ I will assume this rough conception of emergence, as it is a plausible, mainstream commitment and provides a useful starting-point for this discussion.¹²

The combination of these two features – dependence and autonomy - makes emergence mysterious and also somewhat problematic. Dependence on and autonomy from the micro-level aren't mutually exclusive features but, *prima facie*, they stand in tension with one another. As one commentator puts it, understanding emergence in terms of dependence and autonomy *"is like viewing something as both transparent and opaque."*¹³ Most attempts to answer the Interpretation Question involve some attempt to reconcile these two features and dissolve the appearance of tension.

In formulating an account of emergence, a philosopher specifies the exact nature of the dependence and autonomy relation involved in cases of emergence. These relations vary radically across different accounts of and purported cases of emergence. Some accounts of

¹¹ See Wilson, J. (forthcoming), Bedau, M. & Humphreys, P. (2008), Chalmers, D. (2006) for endorsement of this schematic conception of emergence.

¹² I'll focus entirely on cases of purportedly emergent properties, rather than emergent laws, configurations, substances or causes. I will also assume that in most cases a "property" means a property token rather than a type.

¹³ Bedau, M. (2003) pg 156

emergence portray the autonomy characteristic of emergence in metaphysical terms, and others portray that autonomy in merely epistemic or explanatory terms. I will examine these differences in more detail in Chapter 4, but in the meantime we can simply note that different accounts of and purported cases of emergence involve a wide range of different conceptions of this characteristic combination of dependence and autonomy.

The schematic understanding of emergence as involving a combination of dependence and autonomy applies not only to the concept's role in the philosophical literature, but also to its use in science. For instance, scientific practitioners often use "emergent" as a label for natural phenomena (particularly pertaining to properties of complex systems) that resist certain kinds of reductive explanation.¹⁴ In such cases, emergent macro-level properties may be dependent on their micro-level bases in so far as the micro-level properties are properties of the parts that make up the whole that bears the macro-level properties, and autonomous in so far as, for example, they cannot be explained in micro-level terms. The concept of emergence appears in areas as diverse as systems biology and systems chemistry, Artificial Life, fundamental physics and cognitive science.

2.2 A note on the micro-macro distinction

In the philosophical and scientific literature on emergence, cases of emergence are always presented as relative to a distinction between micro-level and macro-level properties. Properties are not emergent *per se*, but instead properties of the macro-level are emergent

¹⁴ For example, see Kelso, J.A. Scott (1995) "It was once thought that single neurons called command neurons might be responsible for generating patterns, a kind of superordinate motor program, but this is not generally so. Complex behavior patterns in neuronal networks are the result of cooperative effects (or as neuroscientists now say, emergent properties) not typically found at the single neuron level." (pg 240-1)

from properties of the micro-level. So in order to understand the concept of emergence, we must understand this distinction between micro- and macro- levels.

There are many different ways to draw a distinction between micro and macro levels, and in the literature on emergence, cases of emergence are presented as relative to a many different micro-macro distinctions. The cases I listed earlier in this section give us some idea of this variety: in case 1) the micro-level properties are the micro-physical properties of the fluid and the macro-level properties are properties of the convection rolls in that fluid, in case 2) the micro-level properties are the properties of cells in Artificial Life worlds and the macro-level properties are properties of the configurations formed by those cells, while in case 3) the micro-level properties are neural properties and the macro-level properties are phenomenal properties. For the purposes of this introduction to the idea of emergence, we can take “micro” to simply mean “more fundamental than macro”. The examples we have considered, however, indicate that “more fundamental” can mean a number of different things.

One way in which micro-level properties can be more fundamental than macro-level properties is for the micro-level properties to be properties of the parts, components or constituents of a whole that bears the macro-level properties. Consider, for example, the difference between the properties of individual members of a population of organisms and the properties of the population itself. I will take the case of the European Herring Gull, *Larus argentatus*, as an example. The population of European Herring Gulls has statistical properties, such as an average wingspan of 144cm, a high relative frequency of living in

shore and estuary habitats and a low relative frequency of living in marsh habitats.¹⁵ And yet any individual European Herring Gull does not have those properties. They have properties such as “lives on an estuary” and “has a wingspan of 140cm”, which in turn are not had by the population of European Herring Gulls. This is one example of a micro-macro distinction drawn in terms of the distinction between the properties of a whole, the population, and the properties of its parts, the individual gulls.

Another way to draw a distinction between micro and macro levels is to claim that the micro-level properties belong to a more fundamental science, as physics stands to chemistry, or neuroscience to psychology.

Jackson and Pettit explore some alternative notions of fundamentality in their discussion of explanatory ecumenism.¹⁶ They describe different ways in which explanations can be fine-grained, including closeness of grain and smallness of grain. When an explanation is close-grained, it gets closer to the causal roots of the phenomenon being explained and so there is typically less of a temporal gap between the explanans and the explanandum. When an explanation is small-grained it gets to the micro-foundations of the phenomenon being explained and so typically the entities that appear in the explanans of a small-grained explanation are smaller in size than those that appear in the explanans of a large-grained explanation. Although Jackson and Pettit focus on explanation rather than fundamentality, their findings indicate some ways in which one level can be more fundamental than another. One level can be more close-grained than another if it is closer to

¹⁵ British Trust for Ornithology. <http://blx1.bto.org/birdfacts/results/bob5920.htm>

¹⁶ Jackson, F. & Pettit, P. (1992)

the causal roots of the features of the higher level, and one level can be more small-grained than another if it is closer to the microphysical level.

Although there are a great many different ways to draw a distinction between micro- and macro – levels, picking out such a distinction isn't just a matter of stipulation. We may choose to focus on a particular micro-macro relationship, such as the relationship between neural and phenomenal properties, rather than another, such as the relationship between physical and chemical properties, but there will be a fact of the matter about whether any two levels do instantiate a “more fundamental than” relation. For the purposes of discussing the concept of emergence as it appears in philosophy and in science, I will leave aside questions about just how many fundamentality relations there are, and what makes for a legitimate micro-macro distinction. We need simply recognize that emergence is always relative to a micro-macro distinction, and that different authors choose to draw that distinction in different ways. In the philosophical and scientific literature on emergence there is no one particular privileged micro-macro distinction, and so the following discussion will not presuppose any particular version of this distinction.

Section 3: The historical development of the idea of emergence

G.H. Lewes first introduced the philosophical use of the term “emergence” in 1874¹⁷ and the concept was taken up and developed by the British Emergentists, a group of British philosophers working in the late 19th and early 20th centuries that included J.S. Mill, C.D Broad and Samuel Alexander as well as Lewes himself. There are some earlier precursors to

¹⁷ Lewes, G.H. (1874)

the concept of emergence. For example, Aristotle claimed that wholes are ontologically prior to their parts and that sometimes wholes are more than the sum of their parts, claims which are similar to those made by contemporary proponents of metaphysical emergentism.¹⁸ But even given these precursors, I will start this historical overview with the British Emergentists as they were the first philosophers to explicitly use the term “emergence”.

C.D. Broad, one of the British Emergentists, developed his account of emergence in response to 19th century debates between vitalists and mechanists. I will examine the vitalist-mechanist debate in more detail in Chapter 4, but in the meantime I will characterize it very roughly as a debate about whether or not certain kinds of explanation were available of certain features thought to be characteristic of living organisms.¹⁹ Mechanists thought that these features, which included metabolism and fermentation, could be explained through appeal to the activities and interactions of the parts of living wholes. Vitalists, on the other hand, thought that these mechanistic attempts at explanation failed, and that in order to explain the characteristic features of living organisms, we had to posit a special kind of substance, a life force or *élan vital*, which endowed living organisms with their special characteristics. C.D. Broad presented the emergent theory as offering a middle ground between these two positions. As Beckermann puts it, “*When C.D. Broad developed his theory of emergence in the early 1920s, his aim was to create room for a third position mediating between these two extremes.*”²⁰

¹⁸ See discussion in Mitchell, S. (forthcoming)

¹⁹ See Beckermann, A. (2000) for discussion of this characterization.

²⁰ Beckermann, A. (2000) pg 1

So what did British Emergentism amount to? Although there were many philosophical differences between the British Emergentists, they were united by the fact that they all appealed to emergence in their attempts to offer a philosophical account of particular relationships between levels in nature. In a definitive historical study of the British Emergentists, Brian McLaughlin argues that their shared position can be characterized in terms of some central features, and I will present those features along with McLaughlin's descriptions of them.²¹

The British Emergentists were materialists. As McLaughlin puts it, "... *British Emergentism maintains that everything is made of matter. There are, for example, no Cartesian souls, or entelechies, vital élan, or the like.*"²² They were also committed to the idea that there are distinct levels of matter, and particular kinds associated with those distinct levels. McLaughlin states, "*According to British Emergentism, there is a hierarchy of levels of organizational complexity of material particles that includes, in ascending order, the strictly physical, the chemical, the biological and the psychological level. There are certain kinds of material substances specific to each level. And the kinds of each level are wholly composed of kinds of lower-levels, ultimately of kinds of elementary material particles.*"²³ The kinds of the different levels have properties particular to the kinds of that level alone, and those kinds have fundamental causal powers. McLaughlin again: "*British Emergentism maintains that some special science kinds from each special science can be wholly composed of types of structures of material particles that endow the kinds in question with fundamental*

²¹ McLaughlin, B. (1992)

²² McLaughlin, B. (1992) pg 49

²³ McLaughlin, B. (1992) pg 50

causal powers.”²⁴ Finally, emergent laws attribute such powers to special science kinds: “*Such a structure will have an emergent causal power as a matter of law, but the law will not be ‘reducible to’ or ‘derivative from’ laws governing lower levels of complexity.*”²⁵

McLaughlin claims that this is an idealized description of the central British Emergentist doctrines, but that it offers a rough representation of the ground that was shared between these thinkers. This shared ground is notable for being naturalistic, as it was based on an attempt to understand scientific results of the time, while also metaphysically non-reductionist, in so far as the British Emergentists took special science kinds to have non-reducible causal powers endowed on them by emergent laws. As such, emergentism lives up to Broad’s attempt to carve out some middle ground between vitalism and mechanism. Unlike vitalism, emergentism doesn’t involve commitment to non-material substances such as *élan vital*, while unlike mechanism, emergentism allows that properties of certain complex entities may have autonomous, irreducible causal powers.

McLaughlin describes the decline of British Emergentism, which he argues was precipitated by scientific developments that rendered emergentism redundant: “*It went wrong for deep empirical reasons. It took many of the greatest scientific achievements of the twentieth century to refute it.*”²⁶ For example, a paradigm case of emergence, according to many British Emergentists, was the case of chemistry. Chemical kinds were thought to have special, irreducible properties endowed upon them by chemical laws. The development of physical models of chemical bonding showed that the emergentist position was wrong;

²⁴ McLaughlin, B. (1992) pg 50-51

²⁵ McLaughlin, B. (1992) pg 51

²⁶ McLaughlin, B. (1992) pg 91

scientists developed physical explanations of chemical bonding. Other scientific developments followed, including the development of molecular biology and genetics, which rendered British Emergentism obsolete.

From the decline of British Emergentism up until the late 20th century, the concept of emergence wasn't at all prevalent either in science or in philosophy. This was a reductionist era, in which the progress of science and of philosophy was measured in terms of the extent to which higher-level phenomena could be explained in terms of more fundamental phenomena and laws. As physicist P.W. Anderson wrote in 1972, "*The reductionist hypothesis may still be a topic for controversy among philosophers, but among the great majority of active scientists I think it is accepted without question.*"²⁷ Since the 1970s, however, a combination of different factors including developments in the sciences of complexity, and philosophical developments such as a renewed interest in questions about consciousness and a special focus on the autonomy of the special sciences, have led to a resurgence of interest in the concept of emergence. In the preface to a recent textbook on the subject, Bedau and Humphreys say

*Thirty years ago emergence was largely ignored in philosophy and science. Its ethos ran counter to the reductionist views of the time, and it seemed to invoke mystical and unexplainable levels of reality. Things have changed. Emergence is now one of the liveliest areas of research in both science and philosophy... The reason for this change is complicated, but it results in part from developments in a number of vigorous and successful research programs within complexity theory, artificial life, physics, psychology, sociology and biology. In parallel, though often driven by independent developments in the philosophy of science and philosophy of mind, philosophers have been developing new conceptual tools for understanding emergent phenomena.*²⁸

²⁷ Anderson, P.W. (1972) pg 393

²⁸ Bedau, M. & Humphreys, P. (2008) pg ix

The concept of emergence is now prevalent across many different areas of philosophy and science. In the next section I will offer a brief survey of different uses of the concept of emergence in contemporary philosophy and science.

Section 4: Emergence in contemporary philosophy and science

In this section I will offer a brief survey of contemporary uses of the concept of emergence in philosophy and in science.

Emergence plays a major role in contemporary metaphysics. Metaphysical uses of emergence typically treat the “autonomy” in “dependence and autonomy” as metaphysical, such that the autonomy characteristic of emergence is metaphysical rather than merely explanatory or epistemological. Emergence is most often appealed to in metaphysics as a way of formulating and defending metaphysical non-reductionism. For example, Merricks defends eliminativism about most composite objects according to which most composite objects are not metaphysically distinct from the atoms that compose them, but he argues that persons are metaphysically distinct from the atoms that compose them in virtue of having emergent properties.²⁹ Alternatively, Chalmers defends a position according to which all facts about the universe are logically determined by the fundamental physical facts, except for the facts about consciousness, and hence that consciousness is emergent.³⁰ Shoemaker also offers a detailed discussion of emergent properties in the development of a non-

²⁹ Merricks, T. (2001) Chapter 4.

³⁰ Chalmers, D. (1996); Chalmers, D. (2006)

reductionist metaphysics of realization.³¹ According to each of these authors emergence tracks metaphysical distinctions, such as the distinction between physical and conscious properties for Chalmers, and between persons and atoms for Merricks.

In philosophy of mind, emergence is also commonly presented as a metaphysical phenomenon, particularly as a way in which properties of conscious experience may be related to other properties such as physical or neuroscientific properties. For example, O'Connor & Wong have argued that the relationship between the mental and the physical is one of “*ontological emergence*”³², where they understand ontologically emergent properties as properties of wholes that are not reducible to relations between properties of the parts of that whole. Kim has explored metaphysical emergentism as a response to the mind-body problem, although he ultimately argues that metaphysical emergentism is self-undermining.³³ In the philosophy of cognitive science, “systems-based” conceptions of emergence have been used to taxonomize characteristic features of dynamical or self-organizing systems. Clark, for example, defines emergence in terms of “collective variables” which he describes as controlling parameters of dynamical systems³⁴, while Port and Van Gelder describe dynamical systems (which, they argue, undergird cognition) as having emergent properties.³⁵

In philosophy of science, Wilson has recently defended an account of emergence understood in terms of “degrees of freedom”, where a degree of freedom is, as she puts it,

³¹ Shoemaker, S. (2007) Chapter 4

³² O'Connor, T. and Wong, H. (2005)

³³ Kim, J. (2006)

³⁴ Clark, A. (1997) Chapter 6.

³⁵ Port & Van Gelder (1995) Chapter 1

*“an independent parameter needed to characterize an entity as being in a state functionally relevant to its law-governed properties and behavior.”*³⁶ Wilson appeals to this notion of emergence in her defense of non-reductive physicalism, which is motivated by a particular concern to recognize the autonomy of the special sciences. Some philosophers, including Teller and Silberstein & McGeever, have argued that fundamental physics reveals cases of metaphysical emergence.³⁷ In the mid-20th century, Hempel and Oppenheim argued for an epistemic account of emergence, according to which emergence marks the limits of our scientific knowledge.³⁸ And finally, Humphreys and Bedau have developed computational accounts of emergence that focus particularly on purported cases of emergence generated by computer simulations and in Artificial Life worlds.³⁹ In these cases, the autonomy characteristic of emergence is understood in computational, rather than metaphysical, terms.

The concept of emergence appears also in many different areas of scientific practice. In Artificial Life, for example, the term is often used for properties of the complex configurations that appear in Artificial Life worlds.⁴⁰ In systems biology and systems chemistry, the term “emergence” is sometimes used as a label for complex phenomena that are particularly surprising or hard to predict from the behavior of their simpler components. Emergence also crops up in areas as diverse as developmental psychology and ecology. For example, a recent article in *Science* describes mussel bed eco-systems as displaying

³⁶ Wilson, J. (2010) pg 281-2

³⁷ Teller, P. (1986); Silberstein, M. & McGeever, P. (1999)

³⁸ Hempel, C. & Oppenheim, P. (1965)

³⁹ Humphreys, P. (2008); Bedau, M. (2003)

⁴⁰ Bonabeau & Theraulaz in Langton, C. (1995)

“emergent” self-organization, while in a *Brain and Behavioral Sciences* article and discussion from 2003, Rumbaugh et al describe “*new forms of behavior that are adaptive*” as “*emergent*”.⁴¹

In addition to its uses in philosophy and science, the idea of emergence has been documented in the popular press, particularly in a number of popular science books of the later 20th and early 21st centuries. Authors including Stephen Johnson⁴² and John Holland⁴³ have written popular books about emergence, focusing particularly on purported cases of emergence generated by complex systems.

In Chapter 4 I will offer a detailed survey of contemporary appeals to the concept of emergence in philosophy and science, but the very brief sketch I have offered here is enough to provide a snapshot of the wide range of different accounts of and purported cases of emergence.

Section 5: Associated Ideas

“Emergence” is one of a group of concepts associated with the possibility that there may be distinct levels in nature. In this section I will give a brief survey of some of these concepts, focusing particularly on *reductionism*, *physicalism* and *supervenience*.

5.1 Reductionism and Non-reductionism

⁴¹ Rumbaugh et al (2003); Koppel et al (2008)

⁴² Johnson, S. (2002)

⁴³ Holland, J. (1999)

“Reductionism” is a term used in many different areas of discourse. The preface to a recent compilation on the subject states, “*Roughly speaking, to reduce is to show that that which is reduced is nothing over and above that which it is reduced to.*”⁴⁴ This is a schematic characterization, but it captures the general idea. For example, we may say that to reduce chemistry to physics, we must show that the kinds, properties and entities of chemistry are nothing more than the kinds, properties and entities of physics. *Reduction* and *reductionism* are common terms, used quite widely in areas outside of philosophy, and exactly how best to move beyond this schematic, general conception of reduction to something more precise is a much-debated topic in philosophy.⁴⁵

In the mid 20th century reduction was commonly understood in terms of laws, following a very influential model of reduction offered by Ernest Nagel.⁴⁶ According to Nagel, “*Reduction ... is the explanation of a theory or a set of experimental laws established in one area of enquiry, by a theory usually though not invariably formulated for some other domain.*”⁴⁷ Nagel argued that in some cases of reduction, which he called *heterogeneous reductions*, generalizations in higher-level sciences are reduced to generalizations in lower-level sciences via principles that establish connections between the kinds of the higher-level sciences and the kinds of the lower-level sciences.⁴⁸ These principles have come to be known as “*bridge laws*”, and Nagel argued that bridge laws not only capture correlations between

⁴⁴ Hohwy & Kallestrup (2008) pg 1

⁴⁵ Nagel, E. (1961); Putnam & Oppenheim (1958); Sklar, L. (1967)

⁴⁶ Nagel, E. (1961)

⁴⁷ Nagel, E. (1961) pg 338

⁴⁸ Nagel, E. (1961) pg 353-354

the kinds of different sciences, but also permit the laws of the higher-level science to be derived from the laws of the lower level science.⁴⁹ Nagel argued that thermodynamics reduces to statistical mechanics in precisely this way.

The idea of reduction played a central role in philosophical discussion of the *unity of science*. Putnam and Oppenheim famously offered a formulation of the idea of the “Unity of Science” and argued for its viability in their 1958 “*Unity of Science as a Working Hypothesis*”. Their conception of the unity of science was made up of two elements: linguistic and explanatory. Linguistic unification is achieved, they argue, when “*all of the terms of science are reduced to the terms of some one discipline*”⁵⁰, and explanatory unification is achieved when “*the laws of science become reduced to the laws of some one discipline*.”⁵¹ They took the combination of explanatory and linguistic unification to be equivalent to (a particular kind of) reduction⁵², and the claim that the unity of science is achievable is, by their lights, a reductionist hypothesis.

The exact details of Putnam and Oppenheim’s model of reduction is not essential to this discussion⁵³, but we should note that they took their central hypothesis (that scientific unity is achievable, if not yet achieved) to indicate that metaphysical emergentism was false.

⁴⁹ Nagel, E. (1961) pg 354. Sklar has argued that bridge laws are identities rather than correlations in Sklar (1967)

⁵⁰ Putnam & Oppenheim (1958) pg 3

⁵¹ Putnam & Oppenheim (1958) pg 4

⁵² Though not equivalent to all conceptions of reduction, as they point out on pg 5 of Putnam & Oppenheim (1958)

⁵³ They adopted the account of reduction developed by Kemeny and Oppenheim. Putnam and Oppenheim(1958) pg 5.

Consider, for example, this passage, which followed in from some discussion of various sources of support for their Unity of Science hypothesis:

*“The consideration just outlined seems to us to constitute an argument against the view that, as objects of a given level combine to form wholes belonging to a higher level, there appear certain new phenomena which are ‘emergent’ in the sense of being forever irreducible to laws governing the phenomena on the levels of the parts.”*⁵⁴

The authors go on to then point out that they are opposed only to this metaphysical claim, and that they acknowledge that there *“are many phenomena which are not reducible by currently available theories pertaining to lower levels”*.⁵⁵

As I mentioned earlier in this chapter, reductionism in philosophy and in science was popular up until the latter half of the 20th century. At this time, a number of philosophers began exploring problems for reductionism of the kind defended by Nagel and by Putnam and Oppenheim. Fodor, for example, published his paper *“Special Sciences”* with the subtitle, *“The Disunity of Science as a Working Hypothesis”* intended to mirror the title of the original Putnam and Oppenheim piece.⁵⁶ Here, Fodor argued for a non-reductionist view of the special sciences, according to which special sciences types, or kinds, don’t reduce, as in a Nagel style reduction, to the types or kinds of other more basic sciences.

If understood as a metaphysical relation, such that the autonomy characteristic of emergent macro-properties is metaphysical, then generally speaking, emergentism is incompatible with reductionism because emergentism posits metaphysically autonomous macro-level properties or entities. If understood as an epistemic relation, emergence can be

⁵⁴Putnam & Oppenheim (1958) pg 15

⁵⁵ Putnam & Oppenheim (1958) pg 15

⁵⁶ Fodor, J. (1974)

compatible with metaphysical reductionism, though of course this depends on the details of both views. As a matter of general use, however, emergence is typically thought of both in philosophy and in science as a “non-reductive” concept, and emergentism as the antithesis of reductionism.

5.2 Physicalism

Physicalism is, schematically speaking, the view that everything is, or is reducible to, the physical. This schema requires us to specify what is meant by “everything”, “is”, “is reducible to” and “physical”, and differences in those specifications make for a wide range of different varieties of physicalism. Jackson and Chalmers have argued, for example, that physicalism amounts to the position that all facts are a priori derivable from the fundamental physical facts.⁵⁷ Ney has argued that physicalism should be understood as a stance, or attitude, rather than a metaphysical position: “*we ought to construe physicalism as nothing more than an attitude, a commitment to form one’s ontology according to whatever the physics of the day says exists.*”⁵⁸ Van Fraassen also argues for the position that physicalism is a stance rather than a thesis.⁵⁹ Regardless of these debates about the correct formulation, however, physicalism has played a significant role in philosophy of mind, philosophy of science and metaphysics of the late 20th and early 21st centuries. So much so that recent

⁵⁷ Chalmers, D. & Jackson, F. (2001)

⁵⁸ Ney, A. (2008) pg 3

⁵⁹ Van Fraassen, B. (2004) Chapter 4

commentators have described physicalism as “*the Weltanschauung of modern analytic philosophy.*”⁶⁰

Part of what has been so attractive for philosophers about the idea of physicalism is that it is seen as a *naturalistic* position. Naturalism is a vague idea, and it is hard to pin down exactly what it means. I won’t attempt to offer a working definition of naturalism here, but will simply say that naturalism is commonly associated with the belief that philosophy is close to and continuous with scientific practice, and is also associated with a set of methodological commitments including the claim that philosophical methods are like scientific methods, and that philosophers should avoid commitment to “supernatural” entities such as gods or ghosts. On this rough characterization of naturalism, we can easily recognize the plausibility of the claim that physicalism may be associated with, and indeed motivated by, naturalism. For philosophers who hold that all sciences in some way are grounded in, if not reduced to, physics, and also hold that philosophy shouldn’t posit anything that cannot be studied by scientists, physicalism would be an attractive position.

Some have argued that physicalism is incompatible with metaphysical emergentism. For example, Jackson and Chalmers’ a priori physicalism is incompatible with Chalmers’ formulation of emergentism.⁶¹ Others have argued that metaphysical emergentism is compatible with physicalism. Crane, for instance, has argued that non-reductive physicalists must endorse emergentism, though he goes on to argue that this makes non-reductive physicalism a problematic position.⁶² The situation is made even more complicated by the

⁶⁰ Loewer and Gillett, quoted in the epigraph of Stoljar, D. (2010)

⁶¹ Chalmers, D. (2006); Chalmers, D. & Jackson, F. (2001)

⁶² Crane, T. (2001)

fact that many philosophers of physics have argued that fundamental physics provides the strongest evidence of metaphysical emergence.⁶³ Regardless of whether or not physicalism in general is compatible with emergentism in general, however, the account of emergence that I will develop in Chapters 3 through 5 is compatible with most conceptions of physicalism.

5.3 Supervenience

Very roughly speaking, set A of properties supervenes on set B of properties if there can be no change in the A-properties without a corresponding change in the B-properties. This idea is often introduced using aesthetic examples, such as, for example, the relationship between the microphysical properties of a statue and that statue's aesthetic properties. The aesthetic properties of the statue may be said to supervene on the physical properties of the statue in so far as there can be no change in the aesthetic properties without a corresponding change in the physical properties. Similarly, we may say that mental properties supervene on physical properties if there is no change in the mental properties without a corresponding change in the physical properties.

The concept of supervenience was widely adopted by philosophers working in philosophy of mind and metaphysics of the late 20th century as tool for expressing the idea that mental and physical properties co-vary with one another. Davidson, for example, stated *"... the position I describe ... is consistent with the view that mental characteristics are in some sense dependent, or supervenient, on physical characteristics. Such supervenience might be taken to mean that there cannot be two mental events alike in all physical respects*

⁶³ Silberstein, M. & McGeever, P. (1999)

but differing in some mental respects..."⁶⁴ However, there are a great many different notions of supervenience in use in contemporary philosophy. For example, some distinguish between global and local supervenience: the former is a relation that obtains between possible worlds, while the latter is a relation that obtains between sets of properties.⁶⁵ Others distinguish between nomic supervenience, which obtains of natural necessity, and logical supervenience, which obtains of logical necessity. Kim also distinguishes between strong and weak supervenience, where weak supervenience obtains between two sets of properties in a *single* possible world, and strong supervenience between two sets of properties across *all* possible worlds.⁶⁶

Supervenience of some kind is clearly necessary for reduction, though there is an ongoing debate about whether or not supervenience is also sufficient for reduction too. Some philosophers argue that supervenience is a reductive relation, while others argue that supervenience merely captures a pattern of property co-variation, and lacks an explanatory element which they take to be essential to reduction.⁶⁷ Either way, most claims about and views on emergence have a supervenience claim built into them, such that emergent macro-level properties supervene in some sense on the relevant micro-level properties. The account of emergence I will defend in Chapters 3 through 5 will involve supervenience, a detail that I will return to discuss in Chapter 4.

⁶⁴ Davidson, D. (2001) pg 214

⁶⁵ Kim, J. (2003)

⁶⁶ Kim, J. (2003)

⁶⁷ Horgan, T.(1993); Kim, J. (2003)

Section 6: Problems for accounts of emergence

Although the concept of emergence is popular and prevalent in philosophy and science, some philosophers have argued that forms of emergentism, particularly metaphysical varieties of emergentism, are problematic and even self-undermining. In this section I will offer a brief survey of some common objections to emergentism, including the problem of downward causation and the objection that emergentism is non-naturalistic and unparsimonious.

6.1 Downward Causation

“Downward causation” is the term for a phenomenon whereby an emergent property has autonomous causal powers to affect other properties or entities at the macro-level and also to affect micro-level properties and entities. Some philosophers have argued that downward causation is a central but also problematic feature of emergence. Kim, for example, argues that downward causation is necessary for emergence but that downward causation is impossible, hence undermining the possibility of emergence.⁶⁸ In this section I will briefly outline the idea of downward causation and why it is thought to be a necessary but also problematic feature of emergence. In Chapter 5 I will return to this topic, and argue that the problem of downward causation is not a problem for my account of emergence.

Given an idea of emergence as a combination of dependence and autonomy, some philosophers, particularly those who favor metaphysical accounts of emergence, have argued that there must be a causal element to emergent autonomy. According to this view, any emergence worthy of the name must involve macro-level properties that are in some sense

⁶⁸ Kim, J. (2006)

causally autonomous. What this means will vary from author to author, but the basic idea is that emergent macro-level properties must be capable of causing effects, and that this capacity is not reducible to the causal capacity of the relevant micro-level properties. If emergent macro-level properties are causally autonomous, however, we face the two problems: a general causal exclusion problem and the particular problem of downward causation.

To illustrate the general causal exclusion problem, consider two sets of properties, P and Q, where P-properties are micro-level properties and Q-properties are macro-level properties, and imagine that all Q-properties are emergent from P-properties. Q^1 is an emergent property, and Q^1 causes the instantiation of Q^2 , another Q-property. However, if all Q properties are emergent from P properties, then Q^2 must itself have a corresponding micro-level P-property, P^2 , from which it emerges. But if Q^2 is emergent from P^2 , then the instantiation of P^2 is sufficient to cause the instantiation of Q^2 , which leaves us with an overdetermination problem: both P^2 and Q^1 are sufficient to cause the instantiation of Q^2 . The overdetermination is systematic; every case of emergence will bring with it an overdetermination problem, which challenges the idea that emergents bring novel causal contributions to the world.⁶⁹

This causal exclusion problem is a problem generated by the possibility of emergent properties being causally autonomous. But the prospect of *downward* causation generates an additional problem for those who endorse the principle of the completeness of physics. I will take downward causation to be the phenomenon whereby an emergent property has

⁶⁹ This is a very rough presentation of the overdetermination problem as presented by Kim in Kim (2006).

autonomous causal powers to affect micro-level properties and entities, such as, for example, a case in which a mental property causes a physical property to be instantiated. I will adopt Crane's rough formulation of the principle of the completeness of physics, according to which "*any physical effect ... is completely fixed, deterministically or indeterministically, by purely physical causes*".⁷⁰ The prospect of downward causation violates the principle of the completeness of physics, because in downward causation, a macro-level property causes an effect at the micro-level, and this cannot be reduced to an instance of micro-micro causation.

Kim argues that the causal exclusion problem undermines the possibility of emergence⁷¹, and Crane argues that downward causation is a problem faced by both emergentists and non-reductive physicalists.⁷² I will save a more detailed discussion of the problem of downward causation for Chapter 5, but in the meantime we should note that both causal exclusion and the possibility of downward causation present a serious problem for proponents of metaphysical emergence.

6.2 Naturalism, Parsimony

Some argue that metaphysical emergentism is non-naturalistic because it requires us to endorse the idea that there are brute facts about macro-level properties and entities, and hence that emergents pose mysteries that science will never explain. As Bedau puts this challenge, "*Strong emergence starts where scientific explanation ends.*"⁷³ This charge

⁷⁰ Crane, T. (2001) pg 18

⁷¹ Kim, J. (2006)

⁷² Crane, T. (2001) pg 19

⁷³ Bedau & Humphreys (2008) pg 159

applies only to strong metaphysical accounts of emergence because epistemic accounts of emergence tend to leave open the option that in our future science we may reduce or explain what it currently emergent. Some have argued that fundamental physics offers evidence for the existence of metaphysical emergence, which provides a counter-challenge to the position that metaphysical emergentism is non-naturalistic.⁷⁴ Even leaving that case aside, however, scientific practice is full of appeals to emergence, made in attempts to taxonomize, account for and describe various different natural phenomena. Given this fact, it seems odd to describe emergentism in general as non-naturalistic, although strong metaphysical emergentism may turn out to be.

Another challenge, related to these claims about naturalism, is the idea that emergentism is ontologically unparsimonious. The emergentist is committed to an extra metaphysical category, the category of emergents, which are not reducible to other metaphysical categories. This again is a challenge applied to metaphysical accounts of emergence, because epistemic emergentism doesn't involve any ontological commitment. Proponents of metaphysical emergence tend to respond to this challenge by arguing that parsimony is a virtue that must be balanced against the other theoretical virtues. If it turns out that emergents appear in the best metaphysical theories, proponents of this position would argue, then, unparsimonious or not, we have reason to believe that emergentism is true.

The view of emergence that I will defend in Chapters 3 through 5 is, I will argue, both naturalistic and ontologically parsimonious.

⁷⁴ Teller, P. (1986), Silberstein & McGeever (1999)

Section 7: Moving on

In Chapter 4 I will offer a more detailed study of the concept of emergence, and of important distinctions such as the distinction between metaphysical and epistemic accounts of emergence. In the meantime I have simply attempted to offer enough background on the concept to set us up for the chapters to come.

Earlier in this chapter, I mentioned that some accounts of emergence portray emergence as a metaphysical relation. According to proponents of metaphysical accounts of emergence, emergence tracks distinctions between metaphysically distinct levels in nature. In the chapters that follow, I will argue that metaphysical accounts of emergence face a problem, called the “collapse problem”. The collapse problem also applies to epistemic accounts of emergence, but I will present collapse objections only to metaphysical accounts of emergence, and in Chapter 4 I will offer an independent argument for the position that the perspectival account of emergence I develop on the basis of my response to the collapse problem is preferable to an epistemic account of emergence. I will return to this point in Section 2 of Chapter 2.

After presenting the collapse problem in Chapter 2, and surveying some failed responses to it, in Chapter 3 I will show that the best way to solve the collapse problem is through an appeal to the connection between emergence and explanation. In Chapter 4 and 5 I will build a positive account of emergence, the *perspectival account of emergence*, and will argue that the perspectival account of emergence doesn’t simply avoid the collapse problem, but that it also offers a unified, naturalistic account of emergence.

Chapter 2: *Collapsing Emergence*

Section 1: Introduction

In this chapter I will reveal a problem that afflicts accounts of emergence. The problem is this: cases of emergence presuppose a distinction between micro-level and macro-level properties. For any purported case of emergence, there are properties that *prima facie* belong to the micro-level, but if they are included in the micro-level, then the purported emergent fails to meet a necessary condition for emergent autonomy. I call these problematic properties *collapse-inducing properties* because when they are included in the micro-level, the purported emergent effectively “collapses”, and yet it seems arbitrary to exclude them. Furthermore, this problem does not depend on the details of any particular account of emergence and so applies quite generally. This is *the problem of collapsing emergence* (or, for short, *the collapse problem*) and I will give an account of it in Section 2. The collapse problem suggests that emergence tracks relationships between arbitrary groups of properties, rather than distinctions between levels of properties in nature.

An intuitive response to the collapse problem is to restrict the micro-level properties in cases of emergence in line with some or another metaphysical distinction, such as the distinction between intrinsic and extrinsic properties. In Section 3 I will survey some of these strategies and argue that they fail. In Chapter 3 I will then move on to offer an alternative solution to the collapse problem, based on the connection between emergence and scientific explanation.

Section 2: The Collapse Problem

An “account” of emergence attempts to characterize (typically by offering necessary and sufficient conditions) what it is for some property to be emergent. In this section, I will present a problem for accounts of emergence (the collapse problem), which will show that a purportedly emergent property fails to meet a necessary condition for emergent autonomy.⁷⁵ This problem is generated by following a recipe, which I will illustrate first by formulating a simple, toy case of a collapse problem and then by formulating collapse problems for three different accounts of emergence: from C.D. Broad, Mark Bedau and David Chalmers. The proponent of the position that there are emergent properties faces the challenge of showing that the recipe for generating collapse problems is illegitimate, a challenge which proves surprisingly hard to meet.

2. 1 A Simple Case

Consider one hundred apples arranged in a circle with a circumference of forty feet. Consider also a person with complete knowledge of the intrinsic properties of each of those one hundred apples.⁷⁶ It is not possible for this person to deduce the circle’s circumference from the intrinsic properties of its constituent apples.

Suppose that a philosopher (let’s call her Em) observes that this is the case and claims that the circle has properties that emerge from the properties of its constituent apples. Em claims that the micro-level properties are the properties of each of the apples, the macro-level properties are properties of the circle, and the emergent macro-level properties are dependent on the micro-level properties in so far as the circle is constituted by the apples, but

⁷⁵ The claims I aim to establish are about emergent properties, but a simple transformation would allow this objection to apply to claims about emergent entities or substances instead.

⁷⁶ We can understand ‘intrinsic’ in the Lewisian sense to mean ‘invariant over duplicates’. Lewis, D. (1986)

nevertheless autonomous in so far as they cannot be deduced from complete knowledge of the micro-level properties. Em presents the failure of deducibility as a necessary condition for emergent autonomy; she takes *having a circumference of forty feet* to be an emergent property of the circle.

However, Em appears to be confused, because whether the circumference of the circle meets the necessary condition for emergence obviously depends on which properties are included in the micro-level. If the micro-level properties include only the intrinsic properties of the apples, then it is impossible to deduce the circumference of the circle from the micro-level properties, and *having a circumference of forty feet* meets the necessary criterion for emergence. If, on the other hand, the micro-level properties include the extrinsic properties of the apples, particularly their relative positions, then there is no failure of deducibility. It seems as if Em has made a mistake in claiming that *having a circumference of forty feet* is an emergent property of the circle, because the moment we broaden the micro-level properties to include extrinsic properties, this property is plainly not emergent. One might argue that Em has not made a mistake because what counts as the micro-level properties in cases of emergence is a matter of stipulation. However, if we don't think that Em has made a mistake, then we take emergence to be an arbitrary, fragile phenomenon that disappears when a wider, and seemingly equally legitimate, group of micro-level properties is taken into consideration.

This is an example of a collapse objection, and it follows a very simple recipe. Starting with the micro-macro distinction used by the proponent of the account of emergence, we find a micro-level property that crosses the barrier that marks emergent autonomy but is also, according to the original micro-macro distinction, a micro-level property. Following

this recipe, the collapse objection generalizes to existing accounts of emergence. In the rest of this section, I will present collapse objections to three different accounts of emergence. I will show that, if we think that Em is wrong to believe that *having a circumference of forty feet* is an emergent property of the circle, then we are committed to thinking that most mainstream philosophical treatments of emergence are wrong for similar reasons.

2. 2 Broad's Emergent Wholes

In the following famous passage, C.D. Broad laid out the commitments of his “*emergent theory*”:

Put in abstract terms the emergent theory asserts that there are certain wholes, composed (say) of constituents A, B, and C in a relation R to each other; that all wholes composed of constituents of the same kind as A, B and C in relations of the same kind as R have certain characteristic properties; that A, B and C are capable of occurring in other kinds of complex where the relation is not of the same kind as R; and that the characteristic properties of the whole R(A, B, C) cannot, even in theory, be deduced from the most complete knowledge of the properties of A, B, and C in isolation or in other wholes which are not of the form R(A, B, C).⁷⁷

For Broad, the micro-level properties are the properties of the parts A, B and C in isolation, while the macro-level properties are the properties of the whole R(A, B, C). Broad claims that in cases of emergence we cannot, even in principle, deduce the properties of the whole R(A, B, C) from complete knowledge of the properties of the parts A, B and C in isolation or in wholes other than R(A, B, C). That is to say, if we knew about the parts, and

⁷⁷ Broad, C.D. (1925), pg 61.

furthermore knew that they were arranged in whole $R(A, B, C)$ ⁷⁸, then we would not be able to deduce the features of the whole. This is presented as a necessary condition for emergence.

There is a lot packed into the claim that the micro-level properties are the properties of A, B and C *in isolation*, and the case of sodium chloride provides an illuminating example. Broad thought that sodium chloride has emergent properties, such as solubility in water, and that its possession of these properties cannot be deduced from full knowledge of the properties of sodium *in isolation* and of chlorine *in isolation*, nor from the properties of other wholes of which they may form parts. From our complete knowledge of the properties of sodium and chlorine *as elements* or in any compound other than sodium chloride, and our knowledge that they are combined together into a compound, we cannot deduce the features of sodium chloride.

To generate a collapse objection to Broad's view, let's focus on part A of the whole $R(A, B, C)$ and let **X**, **Y** and **Z** be the characteristic, and purportedly emergent, features of whole $R(A, B, C)$. Part A has the following dispositional property: *to form a whole with characteristic features X, Y and Z when combined in relation R with parts B and C*. Parts B and C have similar properties.⁷⁹ The emergent features of the whole $R(A, B, C)$ can obviously be deduced from complete knowledge of the features of the parts A, B and C and the knowledge that they are arranged as a whole $R(A, B, C)$, so long as the features of the parts include these dispositional properties. This case of emergence "collapses" when these dispositional properties are included among the micro-level properties.

⁷⁸ Broad doesn't make this element explicit, but a reconstruction of his view suggests that it is necessary because without it Broad's account permits many trivial cases of emergence. For instance, any case in which some parts may be arranged in different ways to form different wholes will generate a case of emergence.

⁷⁹ For the purposes of this discussion, I will assume that A's disposition necessarily manifests itself in response to the stimulus condition of A's being joined in relation R with parts B and C.

Returning to the case of sodium chloride, we can see that these dispositional properties are, by Broad's own standard, micro-level properties. For example, one of the characteristic properties of sodium chloride is its solubility in water. Accordingly, sodium has the following dispositional property: to generate a compound that is soluble in water when combined with chlorine into sodium chloride. Sodium has this dispositional property even if it is never manifested, as would be the case if sodium never came into contact with chlorine. In Broad's terms, this property is a property of sodium "*in isolation*". This claim may be counterintuitive, but by Broad's own lights, the dispositional property is a property of sodium and so a micro-level property.

According to Broad, the failure of deducibility is a necessary condition for emergence. Yet this case shows that whether a macro-level property meets the necessary condition for emergence depends on what counts as a micro-level property. Later, I will consider various kinds of non-arbitrary restrictions on the admissible micro-level properties, such as that micro-level properties not be dispositions. Under this restriction, sodium's disposition to generate a water-soluble compound when combined with chlorine does not count as an admissible micro-level property. But, I will argue, this restriction fails to solve the collapse problem.

2.3 Bedau's Weak Emergents.

Mark Bedau has recently offered an account of "weak emergence" based on cases from Artificial Life, and his approach is typical of proposals rooted in scientific applications of the concept of emergence.⁸⁰ Bedau's account of emergence was developed in a

⁸⁰ Bedau, M. (2003)

philosophical context very different from Broad's, but I will show that Bedau's view is also subject to the collapse objection.

Artificial Life worlds are computer worlds consisting of cells that can be 'dead' (empty) or 'alive' (filled). Filled cells form configurations that change in accordance with simple update rules that specify, given the current configuration of alive and dead cells, what the configuration will be after the next tick of the clock. The state of any given configuration after any interval of time can be derived by simulation; we can simply apply the update rules, over and over again, to determine how the configuration changes with each successive time step. For some configurations, computational shortcuts allow us to find out about the state of the configuration at a given stage without our having to carry out this step-by-step simulation. Other configurations, however, resist such treatment. If we want to find out about the state of *those* configurations after a certain number of generations, we *have* to simulate their development. According to Bedau, such configurations have "*weakly emergent*" properties, and underivability except by simulation is necessary for weak emergence. The macro-level properties, which are the properties of configurations, are dependent on the micro-level properties, the properties of the individual cells, in that the cells constitute the configuration. The macro-level properties are autonomous from the micro-level properties in that there is no computational shortcut by which the macro-level properties can be derived from the micro-level properties.

One example of a weakly emergent property is possessed by the "R-Pentomino" configuration, which appears in Conway's famous Game of Life, an Artificial Life world. This configuration starts off very simply and develops into increasingly flamboyant shapes before settling into a stable state at 1103 generations (the R-Pentomino's "halting point").

Armed with our knowledge of the states of the cells in the first stage of the configuration and our knowledge of the Game of Life's update rule, we can derive the halting point *only* by walking the configuration through each stage in its development. Bedau claims that *having a halting point of 1103 generations* is an emergent property of the R-Pentomino configuration., but Bedau's position is subject to a collapse objection. A filled cell in the Game of Life has the property of *forming a configuration whose development halts at 1103 generations when combined with other cells into an R-Pentomino*.⁸¹ This disposition is a property of the individual filled cell, and by that standard is a micro-property. But if we include this property among the micro-level properties, then we can obviously derive the halting point of the R-Pentomino by means other than simulation.

Of course, any given observer faced with the R-Pentomino for the first time would be unlikely to know that the configuration halts at 1103 generations, and that accordingly any filled cell has the property of *forming a configuration whose development halts at 1103 generations when combined with other cells into an R-Pentomino*. But whether a given observer *knows* that a filled cell has this property is irrelevant to whether it in fact has the property, and having the property is all we need to establish the collapse objection to Bedau.

2.4 Chalmers on emergence as failure of logical necessitation.

According to David Chalmers, almost all of the facts about the world are logically necessitated by the fundamental physical facts, with one very important exception: the case of consciousness.⁸² The facts about consciousness are not *logically* necessitated by the fundamental physical facts, but only *nomologically* necessitated. In worlds with laws like

⁸¹ As in the case of Broad, I will assume that this disposition manifests itself of natural necessity in response to the stimulus condition.

⁸² Chalmers, D. (2006) and (1996)

ours, the same micro-states will give rise to conscious experience, but there are logically possible worlds in which the same micro-states will not give rise to conscious experience. The relationship between the fundamental physical facts and qualitative experience is governed by what Chalmers calls *psychophysical laws*.

Chalmers presents failure of logical necessitation by the fundamental physical facts as a necessary condition for emergence. Unlike Bedau's and Broad's, Chalmers' account of emergence is rooted in contemporary metaphysics, but his view is also subject to a collapse objection. Take a particular individual's conscious experience to have characteristic (and purportedly emergent) features *x*, *y* and *z*, and along with Chalmers take the micro-level properties in question to be the fundamental physical properties of that individual. For the moment let's interpret "*fundamental physical properties*" as the properties of the most fundamental physical entities, and let the "Psycho-Law" be the psychophysical law that governs the relationship between the fundamental physical facts and the relevant features of qualitative experience. Now consider the property of *being governed by the Psycho-Law*. If this is included among the micro-level properties, then the micro-level properties logically, rather than just nomologically, necessitate the macro-level properties. With the inclusion of this micro-level property, the purported case of emergence has collapsed.

One might respond by suggesting that the property of *being governed by the Psycho-Law* does not belong in the micro-level because it is not a candidate for a fundamental physical property. Jackson and Chalmers describe their understanding of 'fundamental' as follows:

Microphysical truths are truths about the fundamental entities and properties of physics, in the language of a completed physics ... We will not engage the issue of

*what counts as “physics”, but will stipulate that if there are any fundamental mental entities or properties, they are not part of physics.*⁸³

The property of *being governed by the Psycho-Law* cannot be the referent of a predicate featured in the language of a completed physics. If we accept the restriction of the micro-level properties to those picked out by predicates in the language of a completed physics, then features *x*, *y*, and *z* meet Chalmers’ criterion for emergence. Note, however, that their doing so *depends* on the fact that this restriction is applied. If the micro-properties include the properties of the most fundamental entities, say, rather than the properties that correspond to the predicates of a completed physics, then the features *x*, *y* and *z* would not meet Chalmers’ criterion for emergence, since *being governed by the Psycho-Law* is a property of the most fundamental entities. Chalmers is similar to Em in this respect; the case of emergence collapses when a wider group of micro-level properties is taken into consideration.

An immediate response to the collapse objection is to think that it involves some kind of cheating. One might suspect that in formulating the collapse problem I have willfully misinterpreted the authors in question, by attributing a broader group of properties to the micro-level than they themselves would recognize. In this paper, I will cut off several attempts to flesh out this criticism by attempting to restrict the relevant group of micro-level properties in cases of emergence. Then, in Chapters 3, 4 and 5 I will offer an account of emergence that not only avoids the collapse objection, but also explains the intuition that the collapse objection involves some kind of cheating.

2.5 The Impact of the Collapse Problem

⁸³ Chalmers, D. and Jackson, F. (2001), pg 316.

Before moving on to consider some responses to the collapse problem, let's take a moment to reflect on what is at stake. In philosophy, emergence has been used to cash out the idea that the universe is composed of metaphysical levels, or layers. For example, the British Emergentists thought that there was a hierarchy of levels of matter, from physics through chemistry and biology, up to consciousness⁸⁴, and some contemporary authors have argued that there is a metaphysical division between consciousness and everything else.⁸⁵ On this kind of picture, there are distinctions between levels of properties in the universe and emergence is the relation that tracks those distinctions. The collapse problem, however, suggests that emergence does no such thing.

A very basic way to think of this problem is in terms of the familiar metaphor of a layer cake.⁸⁶ Think of the universe as a gigantic layer cake, composed of a layer of plain sponge, followed by a layer of cream, followed by chocolate sponge, followed by icing. Emergence is thought to be the relation that tracks the distinctions between those layers, such that the properties of the icing emerge from the properties of the chocolate sponge, which emerge from the properties of the cream and so on. The collapse objection indicates that instead, we have half of the properties of the icing emerging from the properties of the top left quarter of the chocolate sponge and a spoonful of the cream. Unless we can find a solution to the collapse problem, we are left with the result that emergence isn't doing its job; it isn't tracking the right divisions in the cake.

⁸⁴ Broad, C.D. (1925), McLaughlin, B. (1992)

⁸⁵ Chalmers, D. (1996)

⁸⁶ The "layer-cake" metaphor has been used to illustrate many different claims, by philosophers including Sellars. Sellars, W. (1961).

One seemingly attractive response to the collapse problem is to rule out any micro-level properties that in some sense “involve” a property that can only be possessed by the macro-level. If we adopt this restriction, then the collapse-inducing property won’t count as a micro-level property. Obviously this strategy requires a carefully specified notion of “involve”, and the exact details of that notion will determine just how general this solution will be. However, even a general version of this response won’t solve the collapse problem for the simple reason that it permits too much emergence. For example, consider a bundle of pieces of wood, which are arranged into a whole that happens to be a table. Suppose I wonder whether or not features of this whole are emergent, and suppose I use Broad’s criterion for emergence in my investigations. These pieces of wood could be gathered into wholes with any number of different properties: boats, piles of sticks of different shapes and sizes, even art installations. If I cannot include among the micro-level properties any properties that in some general sense “involve” the macro-level, then many properties of the table, such as having four legs and a flat top, will be emergent.⁸⁷ This would not be recognized as a case of emergence by Broad, Bedau, or Chalmers, indicating that this attempt to avoid the collapse problem on their behalf still permits too many cases of emergence.

Alternatively, one might respond to the collapse objection by arguing that there is something special about the three cases I have chosen to present, in so far as each of them involves a failure of deducibility. Perhaps if we consider a case in which the necessary condition for emergence doesn’t involve a failure of deducibility, but rather a metaphysical factor such as a difference in causal powers, then the collapse objection won’t apply. Without

⁸⁷ Some may be worried that artefacts like tables have macro-level properties that are plausibly emergent because they are tied to intentional and qualitative states. To avoid complications, we can stick to simple properties such as “*having a flat top*”.

wishing to spoil the suspense, I will say upfront that this hope is misguided. I will now examine such a case and to show that the collapse objection does indeed apply.

Suppose a philosopher claims that an individual's qualitative experience has properties that emerge from the properties of that individual's brain, in virtue of the properties of the experience having causal powers that aren't exhausted by the causal powers of the brain-properties.⁸⁸ The micro-level properties are the properties of the brain, the macro-level properties are the properties of the experience, and a necessary condition for emergence is that the macro-level properties have causal powers that aren't exhausted by the causal powers of the micro level properties. So, in cases of emergence, one can take all of the causal powers of the micro-level properties and "add" them together and they still wouldn't amount to the causal powers possessed by the macro-level property. The macro-level property has some causal powers over and above the causal powers of the micro-level properties.

It is not at all clear what such "addition" would amount to in the case of causal powers, but we can leave that detail to the proponent of such a view. The main point is that a collapse objection to this position can easily be formulated. Let "E" be a property of this individual's qualitative experience, one of the purportedly emergent properties. Suppose that the micro-level includes the property of *giving rise to qualitative experience having property E*. Then any causal power associated with E is associated with the micro-level properties. Therefore, the purported case of emergence collapses unless the micro-level excludes this property. This case appears to indicate that accounts of emergence that don't impose

⁸⁸ I am assuming that properties have associated causal powers simply for the purposes of exposition, but this assumption could be dispensed with in favour of some other way of accounting for causal capacity.

limitations on the admissible micro-level properties are subject to collapse objections, regardless of whether or not they are based on a failure of deducibility.

So far I have presented the collapse problem as a problem for metaphysical accounts of emergence, which use emergence to track metaphysical distinctions between levels of properties in nature. But the collapse problem also applies to epistemic accounts of emergence, so long as such accounts take emergence to be a non-arbitrary relation. Epistemic accounts of emergence typically treat emergence as tracking the limits of human knowledge, such that when a property is emergent we do not yet know, or cannot explain, the fact that it is instantiated, but that as we build on our knowledge this contingent epistemic gap will be closed. The collapse problem is a problem for any account of emergence that presents emergence as a non-arbitrary relation, where emergence doesn't disappear with the inclusion of some arbitrary property among the micro-level properties. Insofar as epistemic accounts of emergence present emergence as a non-arbitrary relation, the collapse problem applies to them too. However, in Chapter 4 I will provide an additional argument for the position that the perspectival account of emergence that I develop in Chapters 3 through 5 is preferable to an epistemic account, for conceptual reasons. Accordingly, rather than working through the collapse problem as a problem for epistemic accounts of emergence, and considering responses on behalf of the proponent of such an account, I will consider the collapse problem as a problem that poses a serious objection only to metaphysical accounts of emergence.

The collapse problem appears to show that emergence tracks relations between arbitrary groups of properties rather than between distinct levels in nature, but there is an intuitive way to resist this conclusion. If we can find a way to restrict the micro-level properties in line with a metaphysical distinction, then we can avoid the idea that emergence

tracks relations between arbitrary groups of properties. The idea behind this strategy is that we can show that the recipe for generating collapse objections is illegitimate because it involves *the wrong* micro-level properties, where the distinction between “right” and “wrong” micro-level properties is understood in terms of some antecedently recognized metaphysical distinction.

Section 3: Distinctions to Hitch To

Here are my standards for the success or failure of an attempt to solve the collapse problem by imposing a restriction on the micro-level properties. If a proposed restriction lets in collapse-inducing micro-level properties for many mainstream accounts of emergence, and accordingly permits collapses, then it won’t help. If a proposed restriction requires us to rule out many of the properties that one could sensibly wish to include in the micro-level base of an emergent, then it also won’t help, because were we to preclude such properties, we would depart too far from philosophical and scientific practice.

3.1 Intrinsic/Extrinsic

It is tempting to respond to the collapse problem by thinking that collapse-inducing properties are somehow illegitimate; they aren’t genuine, “proper” properties. One way to respect this intuition is to restrict the micro-level properties in cases of emergence to only *intrinsic* properties. The distinction between intrinsic and extrinsic properties is the subject of some controversy, but the basic idea is that a property-bearer has intrinsic properties in virtue of features independent of its environment, and accordingly that intrinsic properties, unlike extrinsic properties, hold across duplicates.⁸⁹ A canonical example is the difference between

⁸⁹ For discussion of the distinction between intrinsic and extrinsic properties, see Lewis, D. (1983); Langton, R. and Lewis, D. (1998); Kim, J. (1982); Weatherson, B. (2001).

mass and weight; the mass of an object is among its intrinsic properties while its weight, which it has in virtue of the force of gravity exerted upon it, is extrinsic. For the purposes of making the appeal to intrinsic properties as attractive as possible, I will assume that there is a distinction between intrinsic and extrinsic properties, that this distinction is metaphysically non-arbitrary, and that we can wield this distinction to some level of accuracy.

In appealing to the distinction between intrinsic and extrinsic properties, the emergentist would claim that emergence is relative to a micro-level composed entirely of intrinsic properties. This appears to be a sensible strategy. Indeed, in the simple case involving apples from Section 2, the collapse-inducing micro-level properties (such as the apples' relative positions) were extrinsic properties, so it might appear that by ruling out extrinsic properties, an emergentist would be able to avoid collapse objections. There are, however, serious problems with this strategy: it is both too restrictive and too permissive.

The appeal to intrinsic properties is too restrictive in so far as there are many cases of extrinsic properties that an emergentist may very reasonably wish to include among the micro-level properties in purported cases of emergence. In many contemporary treatments of emergence, philosophers concentrate on the relationship between neuroscientific (or simply physical) properties and the properties of qualitative experience. These facts about the literature suggest that two paradigm micro-level bases for purported cases of emergence are physical properties and neuroscientific properties. Both physics and neuroscience are, however, full of extrinsic properties. Take *voltage* for example. The property of *being 5 volts* is relative to a ground and accordingly extrinsic, but *being 5 volts* is also a paradigm case of a physical property. *Relative speed* is another example. One particle *having speed of 5 m/s* relative to another is an extrinsic property and another paradigmatic physical property.

Turning to neuroscience, the property of *being a synaptic cleft* is extrinsic; a synaptic cleft is simply a gap across which neurotransmitter molecules are passed from neuron to neuron.⁹⁰

The region of space is a synaptic cleft in virtue of the use that it made of it by the neurotransmitters, which makes being a synaptic cleft an extrinsic property of that region of space. There are many extrinsic properties that an emergentist may reasonably wish to include in the micro-level base of a purported case of emergence, and in most cases these are not collapse-inducing properties.

As well as being too restrictive, the strategy of restricting micro-level properties to just the intrinsic properties is also too permissive. The property that generated the collapse objection to Broad (*having the disposition to form a whole with characteristic features X, Y and Z when combined in relation R with parts B and C*) is plausibly intrinsic⁹¹ and so wouldn't be ruled out by the restriction to intrinsic micro-level properties.

To my claim that the restriction to intrinsic properties is too restrictive, one might respond that for each extrinsic property there may be a corresponding intrinsic property or properties. For example, consider the property of *being a synaptic cleft*. A gap may be a synaptic cleft in virtue of being the site of a certain kind of neurotransmitter activity, but one could argue that this is simply a matter of the site having certain intrinsic properties and other things, such as neurons, having certain intrinsic properties too. If this turns out to be the case, then ruling out extrinsic properties won't be restrictive at all because, so long as we include the right intrinsic properties in the micro-level, we miss nothing by excluding the extrinsic properties.

⁹⁰ Purves, D. et al (2008) pg 96.

⁹¹ The position that all dispositions are intrinsic is popular, but we need not endorse it to take this as a plausibly intrinsic property. See McKittrick, J. (2003) for discussion.

The proponent of such a position faces a dilemma. On the one hand, the promise in the appeal to the distinction between intrinsic and extrinsic properties lies in the idea that it is a metaphysically non-arbitrary distinction. Posit too tight a connection between intrinsic and extrinsic properties and we lose the idea that this is a metaphysical distinction. For instance, someone who claims that each and every extrinsic property is token-identical with an intrinsic property or properties will have trouble defending the claim that there is a genuine metaphysical distinction between extrinsic and intrinsic properties, rather than, say, a conceptual distinction. On the other hand, if the proponent of this objection posits too loose a connection between the relevant intrinsic and extrinsic properties, then they again face the charge of being overly restrictive. For instance, if they agree that not every extrinsic property corresponds to an intrinsic property or properties, then excluding extrinsic properties will still be restrictive. The prospects for stable middle ground are thin, and so this way to save the appeal to the intrinsic/extrinsic distinction looks unpromising. Accordingly I will move on to examine an alternative, the distinction between dispositional and non-dispositional properties.

3.2 Dispositional/Non-Dispositional

As we just saw, the distinction between dispositional and non-dispositional properties appears to be a promising candidate. The basic idea behind this distinction is that what it is to possess a given dispositional property is to have the power to produce certain manifestations in response to certain stimuli. This power exhausts the property and is associated with it as a matter of necessity, while non-dispositional properties lack a stimulus-manifestation essence. Classic examples of dispositional properties include solubility and fragility. The emergentist

could stipulate that the micro-level properties in cases of emergence are exclusively non-dispositional properties, and thereby hope to avoid the collapse problem.

This response assumes a clear distinction between dispositional and non-dispositional properties, and this distinction is a matter of significant controversy.⁹² Some philosophers hold that all properties are dispositional, others hold that no properties are dispositional, and others hold that the question of whether the fundamental properties are dispositional or not is in fact unanswerable.⁹³ Some think that there is a distinction between dispositional and non-dispositional properties, but that this is not a metaphysical distinction.⁹⁴ Some think that there is a metaphysical distinction between dispositional properties and non-dispositional properties, but that every instance of a dispositional property is token-identical with an instance of its categorical basis.⁹⁵ This is a wide and controversial range of options, but for the purposes of this discussion I will adopt the most charitable view possible. I will assume that there is a distinction between dispositional and non-dispositional properties, that we can successfully wield that distinction, and that the distinction is metaphysical rather than, say, conceptual.

This strategy shows promise because dispositional properties often are the properties that generate collapses. In the last section, I mentioned the property that generated the collapse for Broad's case of emergence: *the disposition to form a whole with characteristic features X, Y and Z when combined in relation R with parts B and C*. Once again, however,

⁹² The literature on these questions is vast. See e.g. Blackburn, S. (1990); Molnar, G. (1999); Mumford, S. (1998); Langton, R. (1998); Lewis, D. (2004); Bird, A. (2007); Prior, E. Pargeter, R. & Jackson, F. (1982)

⁹³ Langton, R. (1998), Lewis, D. (2004)

⁹⁴ Blackburn, S. (1990)

⁹⁵ Mumford, S. (1998)

the restriction to non-dispositional properties is both too restrictive and too permissive. For instance, if we turn to the cases of physics and neuroscience, we can see that both are full of dispositional properties. Consider, for instance, the property of *being a working neurotransmitter receptor* (that is, a molecule that is primed to bind a neurotransmitter). This property has the stimulus-response essence of a dispositional property and is a central neuroscientific property. Or, turning to physics, consider the property of *having a mass of 5g*. Some philosophers have suggested that this property is the susceptibility to be accelerated by n cm/s/s by a net force of $5n$ dynes, for any n .⁹⁶ This is again a very central physical property that many people take to be dispositional.⁹⁷ These cases indicate that the restriction to non-dispositional properties is too strong, because, in appealing to this distinction, an emergentist will very seriously restrict the properties they can take to be included in the micro-level. Furthermore, these dispositional properties aren't typically collapse-inducing, which indicates that the dispositional/non-dispositional distinction doesn't track the distinction between collapse-inducing and non-collapse-inducing properties.

The restriction to non-dispositional properties is also too permissive, for there are some non-dispositional properties that generate collapse objections. Consider, for example, the problematic property that generated the collapse objection to Chalmers: *being governed by the Psycho-Law*. Dispositional properties have a stimulus-response essence, but there is no stimulus in the case of the property that generated the collapse objection to Chalmers. Hence,

⁹⁶According to classical physics.

⁹⁷ This claim is somewhat controversial, as endorsing it appears to involve endorsing baseless dispositions. Such controversy depends on the precise conception of the dispositional/categorical distinction in play, however, and so I will leave this question to one side.

this is not a dispositional property even though it generates a collapse. Therefore, the restriction to non-dispositional properties won't rule out all of the problem cases.

3.3 Natural/Non-Natural

According to many philosophers, some properties are privileged in so far as they account for similarity, play a role in causation and the laws of nature and ground other kinds of properties.⁹⁸ These “natural properties” are the categories that really divide up the world, and typically those who endorse the idea that there are such properties follow Lewis⁹⁹ in endorsing the idea that there is a spectrum from perfectly natural to non-natural properties, rather than a binary distinction between natural and non-natural. Many philosophers have appealed to natural properties in response to skeptical problems very similar to the collapse objection. Lewis famously appealed to perfectly natural properties to avoid a similar skeptical problem for his Best System Account of laws and chances¹⁰⁰, and this commitment proved invaluable in solving problems across his philosophical system. It might seem that by restricting the micro-level properties to natural properties, the emergentist can avoid the collapse objection.

There is significant controversy about what naturalness actually amounts to - whether, for example, the natural properties are given to us by *all* of the sciences, or whether they are simply the most fundamental properties.¹⁰¹ The answer to this question will determine just how useful the appeal to naturalness will be for an emergentist. The micro-macro distinction in cases of emergence is a distinction between more and less fundamental levels rather than

⁹⁸ See Lewis, D. (1983); Armstrong, D. (1978); Sider, T. (forthcoming)

⁹⁹ Lewis, D. (1983)

¹⁰⁰ Lewis, D. (1983)

¹⁰¹ Schaffer claims that Lewis and Armstrong vacillate between the two in Schaffer, J. (2004)

between the most fundamental level and some other level. Therefore, if the natural properties appear only in the most fundamental level, then it is impossible for properties to emerge from anything other than that level. Let's again adopt the most charitable interpretation, assuming that metaphysically natural properties are given by all of the sciences, and examine the strategy of restricting the micro-level properties to the natural properties.

Even given these concessions, however, there are some problems with the appeal to naturalness as a way to solve the collapse problem. First of all, naturalness comes in degrees, at least according to Lewis, so we would have to find some point on that spectrum of naturalness at which properties are excluded from the micro-level. If we resort to the idea that properties are non-natural enough to be ruled out of the micro-level only when they generate collapse objections, then this would, of course, be ad hoc. An easy response would be to rule out all but the perfectly natural properties. This move does, however, lead to a very restrictive result. For the perfectly natural properties, even if they are given to us by all of the sciences, presumably form a small group.

The appeal to natural properties also appears not to track the distinction between collapse-inducing and non-collapse-inducing sets of micro-level properties. Consider, for example, the Artificial Life case discussed in Section 3. *Being alive in the Game of Life* and *being dead in the Game of Life* are clearly very important properties of the Game of Life, but they do not have that status outside of the Game of Life, which suggests that naturalness per se may not be enough to account for the inclusion of such properties among the micro-level properties in a case of emergence. *Being alive in the Game of Life* and *being dead in the Game of Life* may play a role rather like the role that natural properties are thought to play,

but only in the very limited context of the Game of Life, whereas metaphysical naturalness is a context-independent matter.

As well as being too restrictive, the appeal to natural properties is also too permissive. Some of the properties that one would need to exclude to avoid collapse objections look fairly natural. Consider, for example, the property of *being the firing C-fiber of a person in great pain*. This property is not perfectly natural, but is not completely gerrymandered either and would certainly have to be excluded from the set of micro-level properties to avoid collapse objections to the claim that properties of qualitative experience emerge from neural properties. Chalmers holds that there are irreducible psychophysical laws that link facts about the physical world to facts about qualitative experience. He doesn't go into much detail about the role of natural properties in such laws, but perfectly natural properties are widely thought to be those that appear in the laws of nature. If this is correct, then Chalmers' view implies that some properties such as *being the firing C-fiber of a person in great pain* will be perfectly natural. Restricting the micro-level properties to only the natural properties still lets in too much.

One might worry that in this discussion we have come perilously close to abandoning a useful and important commitment to naturalness, but in response I would point towards the very particular nature of the task at hand. We are not at the moment engaging with questions about whether or not there are natural properties, whether the universe is fundamentally structured, or whether we can limn that structure. Indeed, I am assuming for the sake of argument that these things are true. The question at hand is whether or not the philosopher who wants to talk about emergence can use the notion of metaphysical naturalness to avoid collapse objections. And the answer to *that* question is no.

3.4 Where do these failures leave us?

So far we have examined three different metaphysical distinctions between kinds of properties, with the aim of establishing whether or not someone who thinks that emergence tracks levels of properties in nature could hitch claims about emergence to one of these distinctions and thereby avoid the threat posed by the collapse objection. I showed that none of these three distinctions tracks the distinction between collapse-inducing and non-collapse-inducing properties. Of course I have considered only three distinctions, but it seems that no metaphysical distinction will do the job.¹⁰²

One response to this result is to take it as an argument against the existence of emergent properties. On this view, it is more plausible to conclude that there are no emergent properties at all than to embrace the idea that emergence can be made to disappear with the inclusion of some arbitrary property in amongst the micro-level properties. Denying emergence is certainly one way to deal with the collapse problem, but in the chapters to come I will explore another, less drastic alternative.

The collapse problem is surprisingly unifying. Accounts of emergence from very different areas of philosophy face the collapse problem. For each purported case of emergence, there may be one particular way to “save” that case and seemingly avoid the collapse problem. For instance, restricting the micro-level properties so as not to include dispositions would avoid the collapse to Bedau’s case of emergence, the halting point of the R-Pentomino, and restricting the micro-level base properties to rule out laws would avoid the collapse for Chalmers, the case of consciousness. At first glance this fact suggests a way to

¹⁰² One approach I haven’t yet considered is Wilson’s account of emergence in terms of eliminations in degrees of freedom, in Wilson, J. (2010). I will quickly note that I do not regard a degree of freedom as a purely metaphysical restriction – a distinction that can be drawn in the absence of considerations about explanation and the goals of investigators. Indeed, I suspect that the degrees of freedom account may end up tracking the explanatory considerations I discuss in Section 4.

avoid the collapse problem; simply find a way to specify the micro-level property *in each case* such that the emergence does not collapse. Yet this attempt to solve the collapse problem only works if we regard emergence as an arbitrary relation, relative to the particular way that we choose to restrict the micro-level properties. Without some principled basis on which to make such restrictions, and without the promise of a general solution to this general problem, rather than an ad-hoc, piecemeal solution, we are committed to a view of emergence as held hostage to a particular way of characterizing the relevant micro-level properties. The solutions I attempted in Section 3 of this chapter were attempts to find a principled basis on which to offer a general solution to the collapse problem by restricting the micro-level properties in cases of emergence, and as we saw, none of these solutions worked.

One might be tempted to think that embracing this result involves embracing the idea that emergence is an uninteresting and arbitrary relation, but in the next Chapter I will argue that this is not the case. According to the solution I will now offer to the collapse problem, emergence tracks the availability and unavailability of certain scientific explanations.

Chapter 3: *An Appeal to Explanation*

Section 1: Introduction

In this chapter, I will defend a novel condition for emergence, *the explanation condition*, and I will argue that this condition provides resources for a solution to the collapse problem. A collapse-inducing property appears to make a purportedly emergent macro-level property non-emergent. I will show that the presence of a collapse-inducing property is not enough to render some property non-emergent. Of course, the property in question may still turn out to be emergent, or non-emergent. I aim simply to show that a collapse objection doesn't settle the matter either way.

I will begin, in Section 2, by outlining the conceptual connection between emergence and explanation, and codifying that connection into the explanation condition for emergence. In Section 3 I will introduce a minimal constraint on scientific explanation, based on the intuition that facts cannot explain themselves. In Section 4 I will combine these resources to offer a solution to the collapse problem. In Section 5 I will explore the motivation for and implications of the appeal to perspectives brought in by the association between emergence and explanation.

Section 2: Emergence and Explanation

Emergence is typically presented as one way in which properties of different levels may be related. For example, in debates in philosophy of mind that focus on the relationship

between neuroscientific properties and the properties of qualitative experience, emergence is an *option*, a way in which the levels may turn out to be related. This *option*, I will now argue, specifically tracks the availability or otherwise of certain kinds of scientific explanation. In this section I will explicate this connection between emergence and explanation, before codifying the connection into the “explanation condition” for emergence.

Many authors are committed to the idea that emergence is associated with the unavailability of explanations. For example, Hempel and Oppenheim offer an account of emergence, and they say of this account: *“If the assertion that life and mind have an emergent status is interpreted in this sense, then its import can be summarized approximately by the statement that no explanation, in terms of micro-structure theories, is available at present for large classes of phenomena studied in biology and psychology.”*¹⁰³ Chalmers explicitly ties emergence to failure of a priori reductive explanation. According to Jackson and Chalmers, failure of a priori reductive explanation is failure of a priori entailment, and the facts about emergent macro-level phenomena are not a priori entailed by the facts about their micro-level bases.¹⁰⁴ Finally, Kim describes the conceptual connection between emergence and explanation: *“the concept of explanation is invoked in the claim that emergent phenomena or properties, unlike those that are merely “resultant”, are not explainable, or reductively explainable, on the basis of their “basal conditions”, the lower-level conditions out of which they emerge.”*¹⁰⁵

¹⁰³ Hempel, C. & Oppenheim, P. (1965) pg 65

¹⁰⁴ Jackson, F. and Chalmers, D. (2001)

¹⁰⁵ Kim, J. (1999) pg 6

Furthermore, in most purported cases of emergence, an attempt at a scientific explanation of the relevant kind has failed. For example, consider the decline of British Emergentism, which I discussed briefly in Chapter 1. The British Emergentists, including C.D. Broad, thought that chemical properties emerged from physical properties. The advent of new physical models of chemical bonding, however, showed them to be wrong. As McLaughlin puts it, “... *quantum mechanical explanations of chemical bonding in terms of electro-magnetism, and various advances this made possible in molecular biology and genetics... made the main doctrines of British Emergentism, so far as the chemical and biological are concerned at least, seem enormously implausible.*”¹⁰⁶ In this case, the British Emergentists were shown to be wrong about certain cases of emergence because of a new explanation of chemical bonding. This connection between emergence and explanation is manifest in the work of authors such as Chalmers, Bedau, Hempel & Oppenheim, Silberstein & McGeever, Kim and Wilson, all of whom present purported cases of emergence as tracking the success and failure of attempts at explanation.¹⁰⁷

These facts about the literature provide evidence for the idea that emergence is conceptually tied to the absence of (at least certain kinds of) explanations. In the rest of this section, I will codify this connection into a condition for emergence, “the explanation condition”. I will argue that, in addition to capturing the conceptual connection between emergence and explanation, the explanation condition will provide resources for a solution to the collapse problem.

¹⁰⁶ McLaughlin, B. (1992) pg 23

¹⁰⁷ Silberstein, M. & McGeever, J. (1999); Kim, J. (1999); Wilson, J. (forthcoming)

I propose a condition for emergence (“the explanation condition”) that captures the connection between emergence and explanation. I will start by simply stating the condition, and then go on to examine and defend its particular features in greater detail.

Where macro-level property “p” is a property of some whole, and parts A, B, C... n are parts of that whole, and where the distinction between whole and parts should be understood loosely, such that any version of the micro- macro distinction discussed in Chapter 1 could correspond to the part-whole distinction, the explanation condition states the following:

A macro-level property p possessed by whole made up of components A, B, C ... n in relation r is emergent iff there is no available explanation of the fact that the following regularity obtains of natural necessity: *Whenever components A, B and C... n are combined in relation r, the resulting whole instantiates property p.*

According to the explanation condition, emergence is relative to a part-whole relation. For example, the property of water-solubility may be emergent relative to the part-whole relation borne by sodium and chlorine to sodium chloride, but not emergent relative to the part-whole relation borne by carbon, hydrogen and oxygen to sucrose. A property is not, according to this condition, emergent *per se*, but is instead emergent relative to a specification of the relevant parts and wholes and their relations. This feature may generate an immediate concern that this position allows for cases of emergence relative to problematic conceptions of the part-whole relation, according to which we can simply generate cases of emergence by picking out the parts in a particular way. I will put this concern to one side for

the moment, but later on in this section and in Chapters 4 and 5 I will argue that this feature does not pose a problem.

To explicate this proposal, I will now specify 1) Why this explanandum? 2) What is meant by “explanation”? 3) What is meant by “available”?

Applying this proposal to the three accounts of emergence examined earlier, from Broad, Bedau and Chalmers, will illustrate exactly what I take the explanandum to be in these particular cases. I will start with Broad’s account applied to the case of the water-solubility of sodium chloride. According to my proposal, if water-solubility is an emergent property, then there is no explanation available of the natural necessity *Whenever sodium and chlorine are combined into sodium chloride, the resulting compound is soluble in water.* Turning to Bedau, if having a halting point of 1103 generations is an emergent property, then there is no explanation available of the natural necessity that *Whenever five individual filled cells are combined into an R-Pentomino, the resulting configuration halts at 1103 generations.* And finally, let’s use the example of properties of qualitative experience when applying my proposal to Chalmers. This is slightly more complicated than the other cases, because Chalmers doesn’t defend any particular instance of the claim that properties of qualitative experience emerge from fundamental physical properties. I will take a person, Jim, who is experiencing red and is in physical state P, as the object of our enquiry. If Jim’s experience of redness is emergent, then there is no explanation available of the natural necessity *Whenever micro-physical entities are combined together into P, Jim experiences red.*

Why, then, should we take this to be the explanandum? First of all, as we can see from the cases above, this proposal fits all the cases we have considered in so far as I have

offered reasonable contenders for the explanandum in each of these cases. Secondly, the concept of emergence picks out one way in which some micro-level properties and one particular macro-level property may be related. The question we are asking when we are investigating whether or not some property is emergent is *why* it is that when a group of micro-level entities are combined into a macro-level entity, that macro-level entity instantiates a certain property. This feature is captured in the general form of the explanandum I have offered.

It is also important to note, as I mentioned before, that the component-whole relation should be understood very loosely, simply in terms of whatever bears the micro and macro-level properties respectively in this particular purported case of emergence. This openness allows the explanation condition to capture a wider range of different cases of emergence than would be permitted by a stricter interpretation of the component-whole distinction. For example, if we adopt an interpretation of the component-whole distinction according to which the components must be object-like pieces that compose a larger object-like whole, rather like a jigsaw puzzle or the pieces of a watch, then we have to exclude as potential emergents a wide variety of purported cases of emergence that don't fit this profile, such as properties of systems that don't decompose neatly into object-like parts, or the virtual objects involved in Artificial Life cases. Rather than impose restrictions on the part-whole relation which would bring with them major restrictions on the kinds of cases of emergence that can obtain, I will leave the part-whole specification and relation quite open, so we can encompass the wide range of different purported cases of emergence and versions of the part-whole relation discussed in Chapter 1.

I haven't yet accounted for the fact that the explanandum involves a natural necessity. Why should we take the explanandum to be the fact that this regularity obtains of natural necessity, rather than as a matter of fact? In response, I would point out that natural necessities are counterfactually stable, non-accidental and therefore more explanatorily potent than other generalizations about relationships between levels in nature, and this partly accounts for their being the objects of interest in investigations between levels in nature. When investigating relations between levels in nature, we are interested in the question of why, *given the laws of nature*, a certain macro-level property is instantiated when a group of micro-level entities are gathered together. This is reflected in the fact that, as I discussed in Chapter 1, some authors argue that supervenience is necessary for emergence.¹⁰⁸ I hold that the right way to conceive of this supervenience relation is as supervenience with respect to the laws of nature, or *nomological supervenience*. Of course, those who don't recognize a distinction between natural and other kinds of necessity wouldn't use an appeal to natural necessities to capture this feature, and so later on in this chapter I will discuss alternative versions of the explanation condition which don't make appeals to a distinction between natural and other kinds of necessity.

I also wish to leave room for the idea that there may be micro-macro laws that govern the relationship between whole and components, as both Broad and Chalmers, among others, have claimed. That there is a micro-macro law would explain the fact that the relation holds, which undermines the conceptual connection between emergence and explanation for which, I will argue, we have substantial evidence. If we take that evidence seriously, as I will argue

¹⁰⁸ See e.g. Crane (2001)

we should, then we should take the relevant explanandum to be a natural necessity rather than a mere generalization.

I have of course helped myself to the notion of natural necessity, and the idea that there is a distinction between logical and natural necessity. I will continue to adopt this framework for ease of exposition, but after presenting my solution to the collapse problem I will argue that neither the explanation condition for emergence nor my solution to the collapse problem depend on these resources. Someone who, for example, thinks that necessity is exhausted by metaphysical necessity could still endorse the explanation condition and my solution to the collapse problem. I will return to this point in Section 4.

I will now move on to address the question of what is meant by “explanation” in “explanation condition”. There are many different sorts of scientific explanation associated with emergence. For instance, Stephan associates emergence with the failure of mechanistic explanation¹⁰⁹, Bonabeau and Therelautz associate emergence with the need for a new model of the relevant phenomenon¹¹⁰, Chalmers and Jackson argue that emergence involves the unavailability of a priori reductive explanation¹¹¹ and Bedau takes underivability except by simulation to be associated with the unavailability of what he calls “micro-level explanation”¹¹². A reasonable way to take this variation into account is to acknowledge that the relevant kind of scientific explanation may vary depending on who is conducting such investigation. For instance, one investigator may be interested in an explanation that involves

¹⁰⁹ Stephan, A. (2006)

¹¹⁰ in Langton, C. (1995)

¹¹¹ Chalmers, D. and Jackson, F. (2001)

¹¹² Bedau, M. (2003)

a particular kind of model, which another may be interested in an explanation that involves decomposing a mechanism.

The driving force behind this idea is recognition of the fact that different sciences, and different ways of practicing those sciences, are associated with different kinds of explanation. Before expanding on this point, it is important to note that a “kind” of scientific explanation is different from a “conception” of scientific explanation. A kind of scientific explanation is simply one of the sorts of explanations offered in science, while a conception of scientific explanation is a characterization of the general role that explanation plays. An example of a conception of scientific explanation is Salmon’s “ontic” conception¹¹³ according to which scientific explanations aim to uncover objective causal relations among events. One may admit many different *kinds* of scientific explanation while maintaining that there is only one accurate *conception* of scientific explanation, and this discussion will not assume a particular conception of scientific explanation. To say, then, that the relevant kind of explanation differs depending on who is conducting the investigation, is not to say that the relevant *conception* will differ.

Many different kinds of explanation are associated with different areas of scientific enquiry. For example, some have argued that neuroscience proceeds through the uncovering of mechanisms¹¹⁴, whereas different approaches to cognitive science are associated with different explanatory styles, including computational, connectionist and dynamicist explanations.¹¹⁵ Idealized models, some argue, play a major role in chemical explanation.¹¹⁶

¹¹³ Salmon, W. (1989) pg 120-121

¹¹⁴ Craver, C. (2007)

¹¹⁵ See Bechtel, W. & Abrahamsen, A. (2002)

¹¹⁶ Weisberg, M. (2011)

Some economists devote themselves to offering explanations of economic phenomena in terms of the rational choices of individuals, while other economists choose to focus on causal, rather than rational-choice, models.¹¹⁷ These cases all give us reason to think that the kind of explanation an investigation aims at will differ across, and often also within, different areas of science, even if we adopt a view according to which there is only one correct *conception* of scientific explanation.

Depending on the kind of investigation at hand, the kind of explanation that the investigation aims at will differ. To explicate this idea, I will introduce a “perspective” as a catchall term for the point of view of a particular observer. Having done so, we can express the current proposal as the idea that the relevant notion of explanation will vary across different perspectives, depending on the kind of investigation that the relevant observers are undertaking. It follows that what it takes for some macro-level property to be emergent from some micro-level properties will differ from perspective to perspective.¹¹⁸

One response to this strategy is to argue that it would be preferable to specify an exact kind of scientific explanation that is associated with emergence, rather than allow for this kind of variation across perspectives that, as I will discuss later, has major metaphysical implications. Sadly, however, this is unsuccessful. The collapse problem affects almost all accounts of emergence, and yet those accounts involve very different notions of scientific explanation. To adopt a connection between emergence and explanation based on only one kind of scientific explanation would not offer a general solution to the collapse problem, whereas the solution that permits variation across perspectives does. Furthermore,

¹¹⁷ Hoover, K. (2001); Becker, G. (1962)

¹¹⁸ In Chapter 5 I will discuss the idea of “perspective” in more detail. For the moment I will simply take this to mean “the point of view of a particular observer”.

understanding emergence as relative to only one kind of explanation presupposes not just a clear definition of scientific explanation, but also an argument for one particular form being privileged. Neither of these is a desirable requirement on an account of emergence, whereas the appeal to variation across perspectives offers a simple, universal solution to the collapse problem, making only minimal assumptions about the nature of scientific explanation. A critic may still worry about the idea that a property's "emergent" status is dependent upon a perspective, but in Chapters 4 and 5 I will argue that a concept of emergence as relative to a perspective not only offers a general solution to the collapse problem, but also offers a unified and naturalistic account of emergence.

Having proposed that when a given macro-level property is emergent, a certain scientific explanation is unavailable, we now face the question of what it is for an explanation to be "unavailable". I propose that what counts as availability may also vary across perspectives, depending on factors such as, for example, the number of people involved in an investigation and their interests. For example, from one perspective, an explanation may be thought to be unavailable if many scientists have made sustained efforts to formulate such an explanation and have been unsuccessful. From another perspective, the standard may be different, such that an explanation is unavailable only if it is impossible for someone in an ideal epistemic position to formulate one. Again, the notion of "perspective" captures the very simple idea that different observers may adopt very different standards, depending on their goals, background knowledge and so on, although multiple observers may well share standards for unavailability as well as for explanation. I will return to a fuller discussion of this point, and the appeal to perspectives in this account of emergence more generally, in Section 5 of this chapter and in Chapters 4 and 5.

Earlier in this section, I argued that a good reason to endorse the explanation condition for emergence is that we already do endorse a conceptual connection between emergence and the unavailability of explanations. I will now offer a further argument for the position that we should endorse the explanation condition, because doing so not only respects the practice of using the concept of emergence to track the availability of certain kinds of explanation, but also offers a solution to the collapse problem.

Section 3: A Minimal Condition for Explanation

In this section I will present a minimal condition for scientific explanation that applies to almost all philosophical accounts of scientific explanation, and is intended to capture the very basic idea that facts cannot explain themselves. In this section, I intend to use the intuition that facts cannot explain themselves to motivate a restriction on whether or not any two sentences can stand in the relationship of explanans to explanandum. The idea behind this strategy is that, although it may be simple to say that facts cannot explain themselves, it is more difficult to say what it is for two sentences to have the same facts as their content. Accordingly I will attempt to formulate a necessary condition for two sentences to stand in the relation of explanans to explanandum, based on how factually close those sentences are to one another, that doesn't involve an account of what it is for two sentences to have the same factual content.¹¹⁹ This minimal condition for explanation, combined with the explanation condition for emergence, will enable me to offer a solution to the collapse problem.

¹¹⁹ For the sake of convenience I will assume that one sentence can explain another, but very little hangs on this locution. These claims could be translated into talk of facts or propositions instead.

As an introduction to the idea that facts cannot explain themselves, consider the fact that bare repetitions of an explanandum statement are universally recognized as failing to explain. Let's take as an example the situation of a cat, Lucy, who has fleas. Lucy's owner wants an explanation of the fact that Lucy has fleas, and so he takes Lucy to the vet. Obviously, if the vet offers "*Lucy has fleas*" as an *explanation* of the fact that Lucy has fleas, then something has gone wrong. Something has gone similarly wrong if the vet offers "*Lucy has the property of having fleas*" as an explanation of the fact that Lucy has fleas. In the second case the vet's attempt at explanation is not quite identical to the explanandum, but still doesn't offer enough to be a contender for an explanation. What Lucy's owner is looking for from the vet is something like "*the brand of flea treatment you have been using is inferior*" or "*Lucy has had contact with a flea-ridden cat and fleas spread*". Of course, there may be many ways in which an attempt at explanation in such a context can go wrong, not in virtue of failing to be an explanation but instead in virtue of being the wrong *sort* of explanation. For example, suppose (improbably!) that in response to the question "*Why does Lucy have fleas?*" the vet starts by describing the Big Bang, the cooling of the Earth, the process of evolution and so on. The vet's response would typically be inappropriate not because it fails to explain why Lucy has fleas, but because it is not the sort of explanation that is relevant in this context. I am interested only in those attempts at explanation that fail because they are in some sense factually *too close* to the explanandum.

Hempel and Oppenheim considered a phenomenon very close to this case, in which an attempt at explanation fails because the explanans is too close to the explanandum: the problem of 'self-explanation'. They consider a case of attempted explanation in which one

premise in the explanans contains a proposition that is replicated in the explanandum, as follows:

Consider the sentences $T_1 = '(x)[P(x) \supset Q(x)]'$, $C_1 = 'R(a, b) \cdot P(a)'$, $E_1 = 'R(a, b)'$... it seems counterintuitive to say that (T_1, C_1) potentially explains E_1 because the occurrence of the component $'R(a, b)'$ C_1 in the sentence E_1 amounts to a partial explanation of the explanandum by itself.¹²⁰

Hempel and Oppenheim consider some attempts to avoid cases that involve the problem of self-explanation, but argue that none of these attempts succeed¹²¹, and leave the problem of self-explanation, in their terms, 'unsolved'. Self-explanation is an example of the phenomenon whereby an attempt at explanation fails because the explanans and explanandum are factually "too close" for one to explain the other.

Another intuitive constraint on explanation, related to the idea that facts cannot explain themselves, is the constraint that the relation between explanans and explanandum should not be symmetric. The basic idea here is that an explanans and explanandum cannot switch roles, because the explanation relation is intuitively *asymmetric*. To introduce this idea, I will briefly describe Hempel and Oppenheim's Deductive-Nomological model of scientific explanation and some counterexamples to it, which illustrate the idea that explanation is asymmetric.

According to Hempel and Oppenheim, for some explanans to offer a D-N explanation of an explanandum, the following must be true:

- i) The explanans must constitute a valid deductive argument for the explanandum

¹²⁰ Hempel, C. & Oppenheim, P. (1948) pg 29

¹²¹ Though they do argue that the purported explanation above shouldn't count as an explanation because the law T is dispensable to the deductive argument. See the rough version of the D-N model which I discuss on pg 70.

- ii) The explanans must include a law of nature as an essential premise of the deductive argument, such that if the premise were excluded, the deduction would not go through.
- iii) The explanans must include empirical content.
- iv) The premises in the explanans must be true.

We should note that the D-N model offered by Hempel and Oppenheim is far more precise. What I have offered here is more a characterization of the explicandum that the D-N model aims at, rather than the model itself, but it is sufficiently detailed for our purposes. There are also a number of qualifications to note for even this rough characterization, including a number of restrictions on what can count as a law of nature, or “lawlike statement”, and the restriction that the explanandum cannot itself be a general or lawlike statement. These details aren’t, however, essential to the current discussion.¹²²

There is a large body of literature responding to the D-N model, including some striking counterexamples. One particular set of counterexamples is interesting for our purposes, those generated by the intuition that in cases of successful explanation the relation between explanans and explanandum is asymmetric. The most famous of these counterexamples is the case of the flagpole and the shadow.¹²³ We can construct a successful D-N explanation in which the length of the shadow cast by a flagpole is explained by the length of the flagpole, combined with the position of the sun and geometrical laws. However, we can also construct an attempt at explanation that meets conditions of the D-N model but is intuitively not an explanation, in which the length of the shadow, combined with the location

¹²² Details on where we can find that discussion in H&O.

¹²³ Originally formulated by Bromberger although the specific flagpole case never appeared in his publications. Cases of a similar sort appear in Bromberger, S. (1966).

of the sun and the laws of geometry, explains the length of the flagpole. This is a counterexample to the D-N model because the length of the shadow intuitively does not explain the height of the flagpole. The difficulties posed by the tower-flagpole case aren't exhausted by the fact that it presents explanation as symmetric, but this is arguably part of the problem with this case.

Responses to these counterexamples vary. Van Fraassen famously described a case in which the length of the shadow does explain the height of (in his case) a tower because the person who built the tower took pains to ensure that it would cast a shadow on a precise spot.¹²⁴ This appeal to context didn't, however, undermine the general idea that the relation between explanans and explanandum is asymmetric. In Van Fraassen's case, it is not simply the length of the shadow that explains the height of the tower, but the length of the shadow combined with some other facts about the builder's intentions, none of which would be explained by the height of the tower. Furthermore, some have argued that what is problematic about the tower-flagpole case is not the fact that it doesn't respect the necessary asymmetry of the relation between explanans and explanandum, but rather that in this particular case the length of the shadow does not explain the height of the flagpole because it's not the kind of feature that could enter into the causal history of the height of the flagpole (except, perhaps, in an outré situation of the sort described by Van Fraassen). As I mentioned before, I don't take the difficulties posed by the tower-flagpole case to be exhausted by the symmetry of the relation between purported explanans and explanandum, but I do take this case to illustrate the intuition that this relation is asymmetric. The restriction on explanation I am about to propose will rule out many cases where the relation between explanans and

¹²⁴ For rather dramatic reasons. See Van Fraassen, B. (1980) pg 134.

explanandum is symmetric, and so these facts about the asymmetry of explanation provide some further motivation for that restriction.

I will now propose a necessary condition for explanation that captures the fact that sometimes two sentences are factually “too close” for one to explain the other, and will also go some way towards capturing the idea that the explanans-explanandum relation is asymmetric.

Let us say that two sentences are “factually equivalent” iff there is no metaphysically possible world in which one sentence is true and the other false.¹²⁵ So, for instance, “*all ravens are black*” and “*all non-black things are non-ravens*” are factually equivalent. “*Lucy has fleas*” and “*Lucy has the property of having fleas*” are factually equivalent, as are “*Lucy has fleas*” and “*Lucy has fleas and $2+2=4$* ”, but “*Lucy has fleas*” and “*Lucy’s owner has been using inferior flea powder*” are *not* factually equivalent. I propose that no explanandum can be explained by a sentence to which it is factually equivalent. This “*factual non-equivalence*” condition is a minimal constraint on explanation, as obviously it takes far more for one sentence to explain an explanandum than for that sentence to be factually non-equivalent to that explanandum. But the factual non-equivalence condition successfully captures the intuition that facts cannot explain themselves.

Although the factual non-equivalence condition applies across almost all accounts of scientific explanation, there is an important class of exceptions, which I will take a moment to identify before setting aside. In any case in which the explanandum is metaphysically necessary, no attempt at explanation will meet the factual non-equivalence condition. If we

¹²⁵ I say “metaphysically possible” rather than “logically possible” because using logical possibility as a criterion fails to rule out many cases in which the purported explanans and explanandum are too close for one to explain the other. For example, “water is able to put out fires” and “ H_2O is able to put out fires” are not far enough apart for either to explain the other, and yet a restriction based on logical possibility wouldn’t rule this case out.

wish to permit explanations of metaphysical necessities, as we should, then we must acknowledge that this class of cases is an exception to the factual non-equivalence condition. Examples of such cases include explanations in mathematics and certain kinds of theological explanation (where the explanandum is a fact about God and as such is metaphysically necessary). For the purposes of this discussion, however, I will put such cases to one side.

It is important to note that the factual non-equivalence condition is a biconditional. Without this feature, any explanation involving deduction, including classic models of scientific explanation such as the D-N model, would not count as scientific explanations. As it stands, D-N explanations involve a conditional, but not a biconditional, such that there are metaphysically possible worlds in which the explanandum is true and yet the premises in the explanans are false. These include possible worlds in which the laws of nature are different, or in which the same result is reached via the same laws from different empirical starting points, such that the “empirical content” in the explanans would be false. The factual non-equivalence condition permits explanations involving deduction, and only rules out as explanations those cases in which there is no metaphysically possible world in which one of the sentences (explanans or explanandum) is true and the other false.

There is another group of seeming counterexamples to the factual non-equivalence condition that we must handle with caution, those cases in which a conceptual truth appears to offer an explanation. For example, consider the case of a scientist in a lab who is faced with a clear, colorless liquid in a test-tube. The scientist knows that this liquid is called “water”, and he wishes to know why the liquid’s meniscus forms a particular curve.¹²⁶ The scientist’s colleague, watching him puzzle over this case, informs the scientist that “*Water is*

¹²⁶ Thanks to Ram Neta for this case.

H_2O ". The scientist knows all about H_2O , and so now has an explanation for the particular shape of the meniscus. It seems in this case that an explanation is provided by a conceptual truth¹²⁷, and hence generates a counterexample to the factual non-equivalence condition for explanation. In response to this case, we should first of all note that if the explanandum is contingent then the factual non-equivalence condition is still met. More importantly, however, I would argue that in this case the conceptual truth provides *access* to an explanation that has otherwise contingent content, and that in other similar cases, a conceptual truth may form *part of* an explanation that is otherwise contingent. Such cases will still meet the factual non-equivalence condition, as only cases in which the explanandum is metaphysically necessary generate counterexamples to the factual non-equivalence condition. These are the cases that I have identified and set aside, and the considerations surveyed in the previous paragraphs indicate that this is a fairly small group, made up most notably of mathematical and theological explanations.

Section 4: A solution to the collapse problem

So far we have established "the explanation condition" for emergence: A macro-level property p possessed by whole made up of components $A, B, C \dots n$ in relation r is emergent iff there is no available explanation of the fact that the following regularity obtains of natural necessity: *Whenever components A, B and $C \dots n$ are combined in relation r , the resulting whole instantiates property p .* We have also established the "factual non-equivalence" condition for explanation: that one sentence cannot explain another factually equivalent to it.

¹²⁷ We need not regard $\text{water} = H_2O$ as a conceptual truth to recognise the force of the problem. Recognising $\text{water} = H_2O$ as a metaphysical truth generates a similar counterexample.

We are now in a position to return to the collapse objection and assess the resources that these findings offer for solving the collapse problem.

A collapse objection is a challenge to the position that some macro-level property is emergent. The presence of a collapse-inducing property supposedly makes the purported emergent no longer meet a necessary condition for emergent autonomy. If a collapse-inducing property is enough to make the relevant macro-level property non-emergent, then according to the explanation condition for emergence, a collapse-inducing property must provide resources for an explanation of the relevant regularity. A “collapse account” is an attempt at explanation based on a collapse-inducing property. In this section I will show that any collapse account fails to meet the factual non-equivalence condition for explanation, and so a collapse-inducing property is not enough to make the relevant macro-level property non-emergent.

I will begin by examining some familiar cases before stating this claim in a general form. In the case of Broad, the explanandum is the fact that the following regularity obtains of natural necessity: *Whenever sodium and chlorine are combined into sodium chloride, the resulting compound is soluble in water.* The collapse objection to the claim that solubility is an emergent property invokes the fact that sodium has the property of *generating a compound that is soluble in water when combined with chlorine into sodium chloride.*¹²⁸ This property provides resources for the following collapse account: *Sodium has (of natural necessity) the property of generating a compound that is soluble in water when combined with chlorine into sodium chloride, and that is why it is naturally necessary that whenever sodium and chlorine are combined into sodium chloride, the resulting compound is soluble in*

¹²⁸ Where, as in the original collapse objection, this disposition manifests itself of natural necessity in response to stimulus conditions.

water. As we can immediately see, the collapse account fails because the purported explanans is factually equivalent to the explanandum.

In the case of Bedau, the explanandum is the fact that the following regularity obtains of natural necessity: *Whenever five filled cells are combined into an R-Pentomino, the resulting configuration halts at 1103 generations*. The property invoked in the collapse objection to Bedau is each individual filled cell's property of *forming a configuration whose development halts at 1103 generations when combined with other cells into an R-Pentomino*.¹²⁹ This property provides resources for the following collapse account: *Each filled cell has the property of generating a configuration that halts at 1103 generations when combined with four other filled cells into an R-Pentomino, and that is why it is naturally necessary that whenever five filled cells are combined into an R-Pentomino, the resulting configuration halts at 1103 generations*. Again, the purported explanans is factually equivalent to the explanandum, and so the collapse account fails.

Finally, in the case of Chalmers, again taking Jim as our example, the explanandum is the fact that the following regularity obtains of natural necessity: *Whenever microphysical entities are combined together into P, Jim experiences red*. In Chalmers' case, the fundamental physical entities have the property of *being governed by the Psycho-Law*, where the Psycho-law is the psychophysical law that governs the relationship between the fundamental physical facts about Jim and Jim's qualitative experience. This property provides resources for the following collapse account: *Each micro-physical entity has the property of generating a whole that experiences redness when combined with other micro-physical entities into configuration P, in virtue of being governed by the Psycho-law, and*

¹²⁹ Again, where this disposition manifests itself of natural necessity in response to stimulus conditions.

that is why it is naturally necessary that whenever micro-physical entities are combined together into P, Jim experiences red. Again, the purported explanans is factually equivalent to the explanandum, and so the collapse account fails.¹³⁰

In each of these cases, we can see that the purported explanans offered in the collapse account is factually equivalent to the relevant explanandum. This shouldn't be surprising, when we consider that the general form of the explanandum is that the following regularity obtains of natural necessity: *Whenever components A, B and C are combined in relation r, the resulting whole instantiates property p.* The general form of the collapse account is *Component A has the property of forming a whole that instantiates property p when combined in relation r with parts B and C.* The presence of a collapse-inducing property doesn't provide resources for an explanation, but merely for a sentence that is factually equivalent to the relevant explanandum. A collapse account is rather like the vet's statement that *"Lucy has the property of having fleas"*. It fails as an explanation in virtue of being factually too close to the explanandum.

On reflection, we can see that this result accords with a common, intuitive response to collapse objections. After presenting collapse objections to Broad, Bedau and Chalmers in Section 2, I mentioned that a standard response to the collapse objection is something along the lines of *"that's cheating"* or *"that's clearly not what the author meant"*. The collapse objections generate dissatisfaction, and the connection between emergence and explanation reveals why this is the case. We use the concept of emergence to track the availability or

¹³⁰ This result depends on how we conceive of the Psycho-law. If it is understood as the general law that governs the relationship between qualitative experience and consciousness, rather than a particular law governing P and Jim's experience of redness, then these two statements are not factually equivalent. In this case, however, the corresponding general statements will be factually equivalent. (*Whenever micro-physical properties are instantiated conscious properties are instantiated, in virtue of being governed by the Psycho-Law, and that is why it is naturally necessary that whenever micro-physical properties are instantiated conscious properties are instantiated.*)

otherwise of explanations, and collapse objections simply bypass that explanatory work. A collapse objection is an attempt to establish “on the cheap” that some property is non-emergent, and so it makes perfect sense to regard collapse objections as illegitimate.

I have now argued for a condition for emergence and a necessary condition for explanation, and have demonstrated that collapse accounts (attempts at explanation based on collapse-inducing properties) don’t meet the necessary condition for explanation. My conclusion (from the condition for emergence) is that a collapse-inducing property is not enough to make a given macro-level property non-emergent. Of course, the failure of the collapse objection does not show that the relevant property *is* emergent. Rather, it shows that a collapse-inducing property cannot decide the matter either way.

Let’s return to the case of Broad to see my approach in action, using the simple case of sodium chloride for illustration. Imagine that we are examining the elements sodium and chlorine and the compound sodium chloride, and investigating whether sodium chloride’s water-solubility is an emergent property. On my view, the collapse-inducing property remains one of the micro-level properties; sodium still has the property *generates a compound that is soluble in water when combined with chlorine into sodium chloride*. However, because of the explanation condition for emergence, sodium’s possessing this property is not enough to make sodium chloride’s water-solubility non-emergent. What *would* be enough would be for an explanation of the relevant sort to become available.¹³¹ In this way the collapse problem is blocked by the explanation condition for emergence.

The results of this section, in addition to providing a solution to the collapse problem, also provide some insight into the failure of the strategies attempted in earlier chapters. In

¹³¹ And, as a historical note, this *was* what happened. Physical explanations of chemical bonding revealed the British Emergentists to be wrong. See McLaughlin, B. (1992)

Chapter 2, I examined various attempts to restrict the micro-level properties in purported cases of emergence (to intrinsic, or non-dispositional, or natural properties). None of these attempts worked, and now we can see why. These metaphysical distinctions failed to pick out what is really important about emergence: the availability or otherwise of scientific explanations. Rather than attempt to eliminate collapse-inducing properties from the micro-level bases of purported emergents, my approach has left them there but rendered them harmless.

In Section 2 I mentioned that, although my solution to the collapse problem presupposes that metaphysical, logical and nomological necessity are distinct, we needn't endorse this view in order to adopt my solution to the collapse problem. This solution can be tweaked to fit a framework in which, for example, modality is exhausted by metaphysical modality. In such a case we could, for example, and depending on the framework, stipulate that both the explanandum and the collapse account are metaphysically contingent but still factually equivalent to each other. Or that both are metaphysical necessities, in which case they would of course be factually equivalent. Similar tweaks can be made for alternative modal frameworks depending on the details of the case at hand, and so this solution to the collapse problem doesn't depend on a particular modal framework.

Section 5: The perspectival move

According to the explanation condition for emergence, if a macro-level property p is emergent, then there is no available explanation of the fact that the following regularity obtains of natural necessity: *Whenever components $A, B, C \dots n$ are combined in relation r , the resulting whole instantiates property p .* I argued in Section 3 that the relevant kind of

explanation and notion of availability will differ across different perspectives, and so what it takes to meet the explanation condition will vary across perspectives. This means that a property's emergence is relative to a perspective; a property that is emergent from one perspective may not be emergent from another.

As I mentioned before, an immediate response to this strategy is to argue that a certain type of scientific explanation is privileged, and hence avoid the appeal to perspectives. I have already argued against this particular move, but I will now consider another strategy of a similar sort. One could simply take all of the different kinds of scientific explanation to which emergence is relative, as a disjunction, and then say that emergence is relative to the disjunction, thereby again avoiding the appeal to perspectives. This may at first seem like a plausible strategy, but my reasons for not taking it are similar to my not arguing for a particular form of explanation being privileged. The basic problem is that, within many different discourses about emergence, it is very important that the absent explanation is of a particular type. For example, in the famous case of the “explanatory gap” in philosophy of mind, to the philosopher looking for an a priori reductive explanation, a mechanistic explanation won't do, and the availability of such an explanation won't render the relevant property non-emergent. I will examine this case in some detail in Chapter 4, but for the meantime we can note that much of the debate over the explanatory gap is conducted over whether mechanistic, neuroscientific explanations of conscious experience are ‘enough’ to close the explanatory gap (in my times, render this phenomenon non-emergent) or whether the only kind of explanation that can do so is an a priori reductive explanation. This is a case in which the absence of a particular *kind* of explanation is important, and not just the absence of *some* kind of explanation.

Earlier, in Section 3 of this Chapter, I argued that different areas of scientific practice are associated with different kinds of explanation. The phenomenon I am pointing towards now is a simple corollary of this commitment. If different areas of scientific practice are associated with different kinds of explanation, and the concept of emergence appears across these different areas, then we should expect the concept of emergence to be associated with the unavailability of different kinds of explanation. And, I have argued in this chapter, this expectation is vindicated. Adopting the strategy of understanding emergence as relative to a disjunct of all of the different kinds of scientific explanation requires us to ignore the differences that genuinely obtain across the different contexts in which the concept of emergence is used.

An objection to the perspectival strategy, related to the idea of leaving open the relevant kind of explanation, is the challenge that the perspectival account of emergence permits deviant cases of emergence. For example, there may be cases of emergence relative to a form of explanation that involves crystal gazing, or the reading of tea leaves, and intuitively such cases ought to be ruled out. In response I would argue that it is indeed important to capture the intuition that there is something wrong with such cases, but that I have resources for doing so without imposing restrictions on the relevant kind of scientific explanation. I deal with this objection thoroughly in Chapters 4 and 5, but in the meantime I will say that my way of handling such objections is to offload the deviance of the case onto the notion of explanation, rather than of emergence. So, such a case may turn out to be a case of emergence in virtue of its explanatory profile, but that case of emergence will be of little scientific or philosophical interest because of the kind of explanation involved. A similar strategy will apply to cases in which emergence is relative to an unrealistic notion of

unavailability, or to an observer with imperfect measuring tools. Chapters 4 and 5 will include a detailed treatment of a range of such cases.

Understanding emergence as relative to a perspective involves understanding emergence as mind-dependent, which has implications for the potential causal role of emergent properties. For many philosophers, much of the interest in the subject of emergence is tied up with the idea that macro-level entities may be causally autonomous in virtue of having emergent properties. Wilson sums up this sentiment in a recent article:

Why care about what emergence is, and whether there is any? To start, many complex entities of our acquaintance (tornados, plants, people and the like) appear to be composed of less complex entities, and to have features which depend, one way or another, on features of their composing entities. Yet such complex entities also appear to be to some extent autonomous, both ontologically and causally, from the entities upon which they depend. Moreover, and more specifically, many “higher-level” entities (particulars, systems, processes) treated by the special sciences appear to be broadly synchronically dependent on “lower-level” (and ultimately fundamental physical) entities. Yet, as is suggested by the associated special science laws, many higher-level entities appear also to be ontologically and causally autonomous, in having features in virtue of which they are distinct from and distinctively causally efficacious relative to the lower-level entities upon which they depend, even taking into account that the latter stand in various aggregative relations. An account of emergence making sense of these appearances would vindicate and illuminate both our experience and the existence and tree-like structure of the special sciences, as treating distinctively real and efficacious higher-level entities and their features.¹³²

According to the proponent of this widely held, attractive position, an account of emergence is worthwhile and interesting in so far as it takes on the question of how macro-level entities can be causally autonomous and yet also dependent on their micro-level bases in a naturalistically respectable way. The findings I have presented pose a serious problem for this conception of emergence. In using emergence to account for the seeming causal autonomy of higher-level entities, we must hold that such entities are causally autonomous in

¹³² Wilson, J. (forthcoming)

virtue of the fact that some of their properties have “emergent” status; this status is a marker of causal autonomy. I have argued, however, that emergence is relative to a perspective. Unless the causal efficacy of higher-level entities is relative to a perspective just like a property’s emergent status, emergence cannot account for the causal efficacy of higher-level entities.

Of course, this finding does not settle the question of whether or not higher-level entities are causally efficacious. Instead, it shows that if higher-level entities are causally efficacious, they are not causally efficacious *in virtue* of some of their properties being emergent. The collapsing emergence findings have broken the link between causation and emergence, though the implications of this finding depend, however, on the view of causation one adopts. A Kantian, a perspectivalist or a subjectivist about causation would probably be able to reconcile the idea of emergence as perspectival with the idea that emergence is causally efficacious, but anyone with a more metaphysical conception of causation is likely to have some difficulty with this result.

These results do not in any way undermine non-reductionism. Indeed, I take it that the motivation for developing a non-reductionist metaphysics, as outlined in Wilson’s quote, is as substantial as ever. But these results undermine the strategy of using the concept of emergence to work out a non-reductionist metaphysics. What the right tools are for this work remains to be seen, whether the best resources for working out a non-reductionist metaphysics will come to us from mereology, causal differences or some other kind of apparatus. In Chapter 5 I will take up this question again.

In this chapter I have offered a solution to the collapse problem, based on the connection between emergence and explanation. In Chapters 4 and 5 I will develop a positive

account of emergence based on this solution to the collapse problem, which I call “*the perspectival account of emergence*”. In Chapter 4 I will develop and argue for an explication of the concept of emergence along these perspectival lines, and in Chapter 5 I will formulate and defend the final version of the positive perspectival account of emergence.

Chapter 4: *An Explication of Emergence*

Section 1: Introduction

There is a general consensus in contemporary philosophy that ongoing debates about emergence are confused and messy. Kim, for example, said in a recent paper “*those discussing emergence, even face to face, more often than not talk past each other. Sometimes one gets the impression that the only thing that the participants share is the word ‘emergence’.*”¹³³ O’Connor refers to emergence as “*a notorious philosophical term of art*”¹³⁴, while Chalmers adds “*The term ‘emergence’ often causes confusion in science and philosophy.*”¹³⁵ In an introduction to a recent collection on the subject, Bedau and Humphreys say, “*the topic of emergence is fascinating and controversial in part because emergence seems to be widespread and yet the very idea of emergence seems opaque, and perhaps even incoherent.*”¹³⁶ Dialogue becomes even more difficult when we turn to scientific uses of the concept of emergence, which are often taken to be radically discontinuous with philosophical approaches. Some claim that the idea of emergence

¹³³ Kim, J. (2006) pg 548

¹³⁴ O’Connor, T. (2006)

¹³⁵ Chalmers. D. (2006) pg 1

¹³⁶ Bedau, M. & Humphreys, P. eds (2008) pg 1

prevalent in science is an entirely different concept from the idea of emergence prevalent in philosophy.

In this chapter, I will attempt to clear up some of this confusion by offering an explication of the concept of emergence that unifies philosophical and scientific uses. Building on the solution to the collapse problem that I offered in Chapter 3, I will argue that the best way to understand the concept of emergence is as the unavailability of a certain kind of scientific explanation for an observer or observers. I will adopt Carnap's criteria for successful explication, and will argue that the explication I offer meets those criteria and fares better against them than alternative explications. I will argue that, in addition to its unificatory appeal, this explication opens up avenues for scientific and philosophical research and dialogue about the nature of emergence. Finally, I will examine two case studies (contemporary debates about the explanatory gap in philosophy of mind and the historical vitalist-mechanist debates) and use them to illustrate the useful work that this explication can perform when applied to particular debates about emergence.

Section 2: Philosophical Approaches to the Interpretation Question

In Chapter 1, I introduced the "Interpretation Question" about emergence: *What is emergence?* In this section, I will examine features of some mainstream philosophical answers to the Interpretation Question.

Given the idea of emergence as a combination of dependence and autonomy, one part of a typical philosophical attempt to answer the Interpretation Question is an account of emergent autonomy. Often such an account will start either with the claim that emergent

autonomy is metaphysical or the claim that emergent autonomy is epistemic. To say that emergent autonomy is metaphysical is to say that emergents aren't "metaphysically exhausted" by their underlying micro-properties, which can be understood in a number of different ways including: a failure of the macro-level to reduce to the micro, the macro-level properties having causal powers that the micro-level properties do not have, or modal differences between the micro-level and macro-level properties. The proponent of an epistemic view, on the other hand, would hold that emergent properties are only epistemically, rather than metaphysically, autonomous from the micro-level. This section will be devoted to developing an understanding of this distinction.

The distinction between metaphysically and epistemically autonomous emergents is tricky because neither of these groups is homogenous. Many different cases fall under the label of metaphysical autonomy and the same is true for epistemic autonomy. In what follows I will examine some examples of metaphysical and epistemic approaches to emergence and reveal that, despite this heterogeneity, there is a reasonable way to draw the distinction. In order to do this, I will examine a paradigm case from each camp: Chalmers' metaphysical view and the epistemic view defended by Hempel and Oppenheim.

According to Chalmers, almost all of the facts about the world are logically necessitated by the micro-level facts, which he takes to be the facts about fundamental physics, with one very important exception: the case of consciousness. The facts about consciousness are not logically necessitated by the micro-level facts, but only nomologically necessitated. That is to say, in worlds with laws like ours, the same microstates will generate conscious experience but there are logically possible worlds in which the same microstates

do not generate conscious experience. This makes conscious experience, according to Chalmers, a case of strong emergence. As he puts it:

*We can think of strongly emergent phenomena as being systematically determined by low-level facts without being deducible from those facts. In philosophical language, they are naturally but not logically supervenient on low-level facts.*¹³⁷

Chalmers' account of strong emergence is an example of a metaphysical account of emergence in so far as it presents emergent autonomy as determined by metaphysical factors, in this case the modal independence of the macro-level from the micro-level.¹³⁸ Not all metaphysical approaches to emergence account for emergent autonomy in terms of modality. Some, for example, hold that macro-level properties are emergent in virtue of causal powers that aren't reducible to the causal powers of the micro-level.¹³⁹ The central characteristic of metaphysical approaches to emergence, however, is that they locate emergent autonomy in some metaphysical, rather than merely epistemic, difference between the micro-level and the macro-level.

Epistemic approaches portray emergent autonomy as merely epistemic rather than metaphysical. For example, consider Hempel and Oppenheim. According to Hempel and Oppenheim, for example:

*"emergence of a characteristic is not an ontological trait inherent in some phenomena; rather it is indicative of the scope of our knowledge at a given time; thus it has no absolute, but a relative character; and what is emergent with respect to the theories available today may lose its emergent status tomorrow"*¹⁴⁰.

¹³⁷ Chalmers, D. (2006)

¹³⁸ Note that, given the collapse problem, emergence is very delicate, and this case is no exception. Whether or not a given macro-level property meets Chalmers' criterion for emergence depends on what micro-level properties get into the micro-level base.

¹³⁹ See O'Connor, T. and Wong, H. (2005)

¹⁴⁰ Hempel, C. & Oppenheim, P. (1965) pg 64

Hempel and Oppenheim offer the following definition of emergence:

The occurrence of a characteristic W in an object w is emergent relative to a theory T , a part relation Pt and a class G of attributes if that occurrence cannot be deduced by means of T from a characterization of the Pt -parts of w with respect to all the attributes in G .¹⁴¹

Hempel and Oppenheim take emergence to be relative to a number of different factors. In Section 1, I mentioned that emergence is always relative to a choice of levels, and this is characterized by Hempel and Oppenheim as a “*parthood relation*”. As they put it, “*Before we can significantly ask whether a characteristic W of an object w is emergent, we shall therefore have to state the intended meaning of the term ‘part of’. This can be done by defining a specific relation Pt and stipulating that all and only those objects which stand in Pt to w count as parts or constituents of w .*”¹⁴² Emergence is also, on their account, relative to a theory, because theories introduce laws and principles that make some features deducible that would not be deducible from a body of knowledge that doesn’t include the theory. For example, Hempel and Oppenheim discuss a case in which electric current flows through a wire connecting a piece of copper and a piece of zinc that are partly immersed in sulphuric acid. They argue that the facts about this flow may not be deducible from basic information about the attributes of copper, zinc and the acid, but may be deducible from a theory that includes these details, plus the principles and general laws of chemistry.¹⁴³ Finally, Hempel and Oppenheim argue that emergence is relative to a class of attributes of the parts.

¹⁴¹ Hempel, C. & Oppenheim, P. (1965) pg 64

¹⁴² Hempel, C. & Oppenheim, P. (1965) pg 62

¹⁴³ Hempel, C. & Oppenheim, P. (1965) pg 63

Social factors play a significant role in Hempel and Oppenheim's account of emergence. In their early, informal characterization of emergence (given above), Hempel and Oppenheim talk of features being relative to "*the theories available today*", and yet in the later, formal characterization of emergence (also given above) there is no mention of time or availability. Emergence *per se* for Hempel and Oppenheim is simply a failure of deducibility. The practice of actually labeling some property as emergent is, however, slightly more complicated, as is revealed in the early informal statement. When we label a property as "emergent", this label is typically relative to the best, or even the most commonly known, theories of the time. This brings an interesting social element into Hempel and Oppenheim's ideas about emergence; emergence *per se* is relative to theory and knowledge, and when we label properties as "emergent" we assume a background of theory and knowledge that, in typical cases, are widely accepted in the relevant social context.

For our purposes, Hempel and Oppenheim's view provides a helpful example of an epistemic account of emergence. They locate emergent autonomy in the scope of human knowledge and the theories we have at hand. As they say of the application of their view to the cases of life and mind, "*If the assertion that life and mind have emergent status is interpreted in this sense, then its import can be summarized approximately by the statement that no explanation, in terms of micro-structure theories, is available at present for large classes of phenomena studied in biology and psychology.*"¹⁴⁴ This last quote also indicates that Hempel and Oppenheim take emergence to be closely connected to the availability of scientific explanation, a connection that will feature prominently in my own account of emergence.

¹⁴⁴ Hempel, C. & Oppenheim, P. (1965) pg 65

These two cases, Chalmers and Hempel-Oppenheim, provide an illustration of the major differences between epistemic and metaphysical accounts of emergence. Epistemic accounts of emergence locate emergent autonomy in the scope of knowledge of an individual observer, a group or a society of scientific practitioners. According to proponents of epistemic accounts of emergence, a macro-level property is emergent relative to a body of knowledge, which makes emergence epistemic, though whether or not a property is emergent *relative to* a body of knowledge may be an objective matter. Metaphysical accounts, on the other hand, locate emergent autonomy in the mind-independent world such that emergence is not relative to a body of knowledge, but exists per se in virtue of metaphysical relations between micro-level and macro-level properties.¹⁴⁵

Section 3: The Concept of Emergence in Scientific Practice

In this section, I will examine the uses to which the concept of emergence is put in purely scientific, rather than philosophical or interdisciplinary contexts. Interest in scientific uses of the notion of emergence has grown over the last few decades as improvements in computational modeling have led to developments in “systems sciences”, which are roughly those areas of scientific practice that focus on complex systems and proceed by building models of those systems. Much scientific dialogue about emergence appears in the context of systems sciences, but the idea of emergence also appears in other areas such as quantum mechanics, developmental psychology and Artificial Life.

¹⁴⁵ There are some other kinds of philosophical accounts of emergence, including computational views, which locate emergent autonomy in the computational irreducibility of the macro-level from the micro-level, but most philosophical accounts tend to fall into the category of metaphysical or epistemic.

The division between philosophical and scientific treatments of emergence isn't entirely clear. Some scientists make philosophical claims about emergence when they employ the concept, and indeed some of the areas in which the idea of emergence features are areas in which scientists and philosophers interact, such as cognitive science and the interdisciplinary cluster around systems science. In this section I will try to avoid blurring the boundaries by examining pieces from scientific rather than philosophical journals and focusing on authors who self-identify as scientific practitioners rather than as philosophers.

Many authors have claimed that scientific and philosophical uses of the concept of emergence are discontinuous, some going so far as to claim that there are in fact two separate concepts of emergence, one corresponding to the concept prevalent in science and one corresponding to the concept prevalent in philosophy. For example, Chalmers draws a distinction between two different concepts, strong emergence and weak emergence, and claims that *"Strong emergence is the notion of emergence that is most common in philosophical discussions of emergence"* while *"Weak emergence is the notion of emergence that is most common in recent scientific discussions of emergence..."*¹⁴⁶ Stephan says of the concept of emergence *"People associate with this concept the hope to adequately classify hard philosophical problems such as the problem of phenomenal qualities and they associate with it the hope to characterize adequately behaviors of artificial systems, which are not explicitly programmed, but rise from self-organizing processes. A closer look, however, reveals that we have no concept of emergence that is apt for both issues."*¹⁴⁷ This brief survey of scientific uses of the concept of emergence will indicate that this is not in fact the case,

¹⁴⁶ Chalmers, D. (2006) pg 1-2

¹⁴⁷ Stephan, A. (2006) pg 485

and that scientific ideas about emergence are continuous with philosophical ideas about emergence.

Systems biology and systems chemistry are areas in which emergence is a fairly prevalent concept. In a recent article on systems chemistry from the journal *Nature*, Nitschke describes systems chemistry as “*the study of complex systems, or networks, of molecules*” and directly addresses questions about the nature of emergence: “*Emergence occurs when a complex system exhibits properties that can’t be predicted by considering its subcomponents in isolation... Predictability is subjective, however... Perhaps a better definition of ‘emergent’ is ‘interesting and counter-intuitive’ which highlights the necessary subjectivity and ties the elucidation of emergent phenomena directly to the scientist’s role in unraveling the predictive rules that underlie such phenomena.*”¹⁴⁸ For this practitioner, emergence is tied to unpredictability.

In another recent piece on systems chemistry, Balazs and Epstein claim that “*The concept of emergence in the physical and biological sciences is an elusive one*” but go on to offer a definition “*The term refers to phenomena in which the complexity of structures or behaviors in systems with many interacting components exceeds that predicted from the knowledge of the individual components and the forces between them.*”¹⁴⁹ Again we can note the appeal to unpredictability, in this case the unpredictability of the *complexity* of certain systems.

¹⁴⁸ Nitschke (2009) pg 737

¹⁴⁹ Balazs, A. and Epstein, I. (2009) pg 1632

Another area of scientific practice in which the concept of emergence is prevalent is Artificial Life. A journal serving the Artificial Life community describes Artificial Life as “*a new discipline that investigates the scientific, engineering, philosophical, and social issues involved in our rapidly increasing technological ability to synthesize life-like behaviors from scratch in computers, machines, molecules, and other alternative media.*”¹⁵⁰ Assad and Packard claim that, “*since its inception, the field of Artificial Life has consistently referred to the property of ‘emergent’ phenomena as one of its distinguishing features*”¹⁵¹ but go on to express concern about the absence of a definition of emergence. They propose a graded scale of emergence from strong to weak, where each place on that scale represents a different form of failure of deducibility. In the weakest cases, the facts about the emergent features are deducible in hindsight after observing the relevant phenomenon but not before (presumably because the observation provides new information about the micro-level properties, though this isn’t specified by the authors), while in the strongest cases the emergent features of the phenomenon are impossible to deduce from a full specification of the relevant micro-level properties. This is under-specified, and Assad and Packard are unclear about whether they see this scale as a definition of emergence, a heuristic for picking out cases of emergence or something else. For our purposes, however, we can again note the appeal to unpredictability, and also to failure of deducibility. These themes are evident in other writings on Artificial Life. Bonabeau and Theraulaz, for instance, say: “*emergence is generally defined as a process through which entirely new behaviors appear, whose properties cannot be derived from a given model of how the system behaves, so that another model has to be built in order*

¹⁵⁰ Mission statement from the journal *Artificial Life*

¹⁵¹ Assad, A. & Packard, N. (1992) pg 231

to deal with these new behaviors.”¹⁵² This extract is interesting because the phrase “*entirely new behaviors*” has metaphysical connotations, but the newness of these behaviors is understood in terms of a failure of derivability within a given model.

The concept of emergence is currently prevalent in fundamental physics. Fundamental physical notions of emergence have received much philosophical attention, most of which treats such emergence as metaphysical. For example, Silberstein and McGeever argue that “*fundamental physics provides us with the strongest empirical evidence for ontological emergence*”¹⁵³, where ontological emergence is understood as “*features of systems or wholes that possess causal capacities not reducible to any of the intrinsic causal capacities of the parts nor to any of the (reducible) relations between the parts.*”¹⁵⁴ Because purported cases of emergence in fundamental physics have received so much philosophical attention I will leave them out of this section, pausing only to note that these typically fall into the category of philosophically detailed metaphysical accounts of emergence.

Finally, a historical treatment of the concept of non-reductionism reveals both metaphysical and explanatory ideas about emergence. In an article called *More is Different*, published in the journal *Science* in 1972, Nobel-laureate physicist P.W. Anderson presented some challenges to what he called the “reductionist hypothesis”. As he puts it,

“The behavior of large and complex aggregates of elementary particles, it turns out, is not to be understood in terms of a simple extrapolation of the properties of a few particles. Instead, at each level of complexity entirely new properties appear, and the

¹⁵² In Langton, C. (1995) pg 306

¹⁵³ Silberstein, M. & McGeever, P. (1999) pg 184

¹⁵⁴ Silberstein, M. & McGeever, P. (1999)pg 182

understanding of the new behaviors requires research which I think is as fundamental in its nature as any other.”¹⁵⁵

Although Anderson does not use the term “emergence”, it is clear that he is talking about complex phenomena that resist reductive explanation. Anderson gestures towards both a metaphysical claim and a claim about the explanatory status of these phenomena. The metaphysical claim is that there are new properties and entities at higher level of complexity, while the explanatory claim is that complex systems cannot be explained in terms of the simpler parts that compose them.

In this section I have very briefly examined some extracts from practitioners working in systems chemistry, Artificial Life and physics. In each extract, some claims about emergence were made, many of them vague and schematic by philosophical standards. Despite this vagueness, however, the following common features were evident: emergence was tied to unpredictability, failures of derivability, and resistance to certain kinds of explanation. Some metaphysical underpinnings for such features were gestured towards, but not explicitly addressed.

In passing, it is important to note that, when talking about the unpredictability and underivability of facts about emergence, practitioners typically mean unpredictability or underivability *from* a restricted set of facts. Practitioners will tacitly understand at least roughly what set of facts is at issue. Unpredictability or underivability is typically understood as unpredictability or underivability from only the facts about the micro-level, and excluding past experience of similar systems, configurations, compounds and so on. This qualification is necessary if attributions of emergence in scientific contexts are to remain consistent.

¹⁵⁵ Anderson, P.W. (1972) pg 393

As I mentioned at the start of this discussion, the position that there are at least two radically different concepts of emergence, one of which roughly corresponds to philosophical use and one to scientific use, is fairly popular. The extracts we have examined from scientific practitioners suggest, however, that this position is false. Features such as unpredictability, resistance to certain kinds of explanation and difference in metaphysical status are features of philosophical as well as scientific treatments of emergence, which indicates some level of commonality between those conceptions.

Section 4: The Need For a New Approach

The philosophical literature on emergence is notoriously messy. Proponents of metaphysical and epistemic accounts of emergence will often take emergence to obviously be a metaphysical or an epistemic phenomenon, and the approach of *starting* with a particular idea of emergence as necessarily epistemic or necessarily metaphysical has led to a situation in which there is little common ground in philosophical debates about emergence. Wilson's pessimism about this state of affairs is representative: "*Though a thousand flowers may fruitfully bloom, this much diversity is unuseful for the purposes of illuminating the structure of natural reality. Different accounts often disagree on whether an entity is emergent, and when they agree there is often no clear basis for this agreement.*"¹⁵⁶ Scientific uses of the concept of emergence generate even more confusion. As was emphasized in Section 3, many authors take the scientific and philosophical concepts of emergence to be radically discontinuous, though the extracts surveyed in Section 3 indicate that this is not the case.

¹⁵⁶ Wilson, J. (forthcoming) pg 3

The current situation is that philosophers with different views on emergence typically understand “emergence” as a term of art, different in the hands of each person who uses it. This makes philosophical dialogue about emergence troublesome and dialogue between philosophers and scientific practitioners even more so. If there really are numerous different concepts of emergence being applied to different kinds of phenomena, then this diversity isn’t a problem so long as individuals are explicit about the concept they are deploying and deploy it appropriately. In what follows, however, I will offer a rational reconstruction of the concept of emergence that will unify philosophical and scientific uses. In doing so I will show that one concept can serve the majority of purposes to which the concept “emergence” is typically put, and that furthermore using this unified concept will open up interesting avenues for philosophical and scientific dialogue about and research into emergence. This rational reconstruction will be based on my solution to the collapse problem, so in addition to its conceptual benefits this explication of emergence will also avoid the collapse problem.

The basic idea of a rational reconstruction is to take a vague but prevalent concept from a domain of discourse and to offer a precisification of that concept based on features of its use.¹⁵⁷ Examples of domains of discourse include scientific practice, medicine, law, religion and mathematics. Although rational reconstruction aims at accuracy to original use, it isn’t determined by original use, and allows for revisions to the concept in question. Carnap gives a rational reconstruction of probability in his *Logical Foundations of Probability*, and at the start of that text he offers an account of rational reconstruction, which he refers to as “*explication*”. (Henceforth I will treat “rational reconstruction” and “explication” as interchangeable.) As Carnap puts it: “*The task of explication consists in*

¹⁵⁷ The notion of explication was historically associated with positivist ideas about the proper role for philosophy of science, which I don’t share. I am merely using this as a method for clarifying a concept.

transforming a more or less inexact concept into an exact one or, rather, in replacing the first by the second. We call the given concept ... the explicandum and the exact concept proposed to take the place of the first ... the explicatum."¹⁵⁸ He goes on to offer a characterization of an adequate explication, according to which an explication is better insofar as it more fully satisfies the following desiderata:

1. Similarity. *"The explicatum is to be similar to the explicandum in such a way that, in most cases in which the explicandum has so far been used, the explicatum can be used; however, close similarity is not required, and considerable differences are permitted."*
2. Exactness. *"The characterization of the explicatum, that is, the rules of its use... is to be given in an exact form, so as to introduce the explicatum into a well-connected system of scientific concepts."*
3. Fruitfulness. *"The explicatum must be a fruitful concept, that is, useful for the formulation of many universal statements..."*
4. Simplicity. *"The explicatum should be as simple as possible; this means as simple as the more important requirements 1, 2 and 3 permit."*¹⁵⁹

In the next section I will offer a rational reconstruction of the concept of emergence and will adopt Carnap's criteria for successful explication as my own standards for success.¹⁶⁰ The concept of emergence is, however, a tough target for a rational reconstruction because it is

¹⁵⁸ Carnap, R. (1950) pg 3

¹⁵⁹ All quotes taken from pg 7 of Carnap, R. (1950)

¹⁶⁰ For further discussion of what makes for an adequate explication, see Quine, W.V.O. (1960) and Hanna, J.F. (1968).

used in philosophy as well as in scientific practice, and so the explicandum has already received some substantial philosophical attention. I will attempt to work around this feature by simply taking the explicandum to be the concept of emergence as it is used across both scientific practice and philosophy, and to offer an account that unifies those uses as far as possible.¹⁶¹

The task of rationally reconstructing a concept is different from that of offering a full philosophical account of the phenomenon, if there is one, to which that concept applies. In this piece I will offer a rational reconstruction of the concept of emergence, but that reconstruction will leave open serious philosophical questions about the nature of emergence. One could formulate a successful rational reconstruction of the concept of emergence even if there were no cases of emergence in the world, only people using the concept. A philosophical account of emergence will answer questions, particularly metaphysical questions, left open by a rational reconstruction. For an example of this distinction, consider the difference between a reconstruction of the concept of causation and a full philosophical account of causation. A reconstruction of the concept of causation may include features such as transitivity and time-asymmetry. The philosophical account of causation then has to make room for such features, showing that they have a certain place in a metaphysical landscape, for example, or perhaps explaining them away as features of perspectives. In this paper, I will merely offer a reconstruction of the concept of emergence, leaving the completion of the development of my full account of emergence for Chapter 5.

¹⁶¹ I will not include everyday uses of the term “emerge” as in “*he emerged from the shop*” or “*the train is emerging from the tunnel*”.

Section 5: An Explication of Emergence

I offer the following explication of the concept of emergence:

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

In the rest of this section I will examine some of the central features of EM, before moving on to assess EM against Carnap's criteria for successful explication in Section 8.

First of all, I should point out that "components" should be understood very loosely. I simply take the term "component" to be a placeholder for "bearer of the micro-level properties" and the term "whole" as a placeholder for "bearer of the macro-level properties", where the distinction between micro and macro can differ from case to case along the lines described in Chapter 1. This means that emergence is relative to a component-whole relation, as well as to a particular specification of the components.

According to EM, a property isn't emergent *per se* but only emergent relative to an observer, among other factors. Although the conception of emergence given in EM refers to only one observer, EM allows for the possibility that a given property may be emergent for many observers, or even all, while another may be emergent for only one, or a few observers.

I take "explanation" to be a scientific explanation, but the particular kind of explanation in question can vary from observer to observer, depending on a range of factors

including the area of science in which that observer is working, the tools available to them and so on. I have argued in detail elsewhere that this is the correct form for the explanandum and won't repeat that argument here¹⁶², but let me emphasize that the explanandum is a natural necessity rather than a simple generalization. This builds a nomological supervenience relation into this explication of emergence, such that emergent macro-level properties supervene of natural necessity on their micro-level bases. The relevant notion of "availability" also depends upon the relevant observer. One observer may take an explanation to be unavailable if a group of top scientists have worked on the question for a long time and have not succeeded in formulating such an explanation. A different observer may have a different standard, such that an explanation is unavailable only if it is impossible to formulate one.

One striking feature of EM is that it portrays emergence as extremely prevalent. For instance, if the relevant observer has very limited scientific knowledge, then for that observer there may be very many more cases of emergence than for an observer with more extensive scientific knowledge. If we endorse EM, we endorse the idea that each of these cases is a genuine case of emergence, relative to the relevant observer. Some might take this to be an unfortunate feature, but I would argue that it is unproblematic. The prevalence of emergence would be problematic if we had to treat all such cases as equal, but EM provides resources to address questions about why some cases of emergence are more interesting than others. For instance, we can focus on differences between the capacities, tools and knowledge of different observers and on differences between the different sets of micro-level properties. If some macro-level property is emergent simply because the observer has a misguided

¹⁶² See Chapter 3

conception of scientific explanation, then this is an uninteresting case of emergence. A more interesting case could involve the fact that a central, important *kind* of scientific explanation has not yet been recognized. Or alternatively, another kind of interesting case could involve the fact that an explanation of a certain kind is, although recognized as a kind of explanation, unavailable in this particular case, and that furthermore this is true for all observers in the contemporary scientific community. Given that EM doesn't treat all cases of emergence as homogeneous, the prevalence of emergence isn't a problem for this reconstruction. There may be many cases of emergence, but EM allows for great variation among the different cases.

This last point is an immensely important element of the conception of emergence provided in EM. EM offers a conception of emergence as a prevalent and varied phenomenon. In endorsing the concept of emergence offered in EM, we endorse a shift in dialogue about emergence. We ask not just *whether* some property is emergent, but also *why* and *for whom* it is emergent. On this conception of emergence, there are many different cases of emergence that obtain for many different ranges of observers and for many different reasons.

A fan of metaphysical accounts of emergence might argue that EM is an inadequate explication of emergence because metaphysical accounts of emergence typically don't present emergence as relative to an observer. Instead, metaphysical accounts of emergence typically present emergent macro-level properties as emergent *for everyone*, which makes relativization to an observer redundant. In response I would point out that EM portrays emergence as relative to an observer, but this doesn't mean that each case of emergence is emergent *for only one observer*. Some cases of emergence may be emergent relative to many

observers, and others relative to only a few. Some cases may be emergent for all possible observers, and nothing in EM rules out such an eventuality. EM makes room for cases of emergence in which the emergence obtains for metaphysical reasons (in which case the relevant property would be emergent for all observers for all time) but also for cases in which the emergence obtains merely for epistemic reasons. I will return to this point when assessing objections to EM.

Someone may wonder why EM does not include any reference to unpredictability, or to a failure of deducibility, given that these are such prevalent features across different accounts of emergence. In response I would point out that in other writing I have argued that accounts of emergence which take such features to be necessary for emergence face a serious problem, the collapse problem, the only viable solution to which involves understanding emergence as intimately connected to explanation. More importantly, however, EM can make room for conceptions of emergence as sometimes (though not necessarily) involving a failure of deducibility, or the unpredictability of facts about emergent macro-level properties from facts about micro-level properties, depending on the relevant conception of explanation. Some kinds of explanation involve deduction and if such explanations are unavailable, then the case of emergence will involve a failure of deducibility. EM makes room for such cases, but avoids the collapse problem that afflicts accounts of emergence that take failures of deducibility or predictability to be necessary for emergence.

Those familiar with the philosophical literature on the explanatory gap will recognize the appeal to the unavailability of an explanation as an element of the explanatory gap. There are many different formulations of the explanatory gap, based around different ideas about explanation and about the target phenomenon, but the rough idea is that facts about

qualitative experience cannot, even in principle, be deduced from physical facts. In Section 9 of this chapter I will examine the case of the explanatory gap in some detail, but for the moment we can note that if there is an explanatory gap, then according to EM it presents a particularly prevalent and intriguing case of emergence.

Section 6: Considering Carnap's Criteria

In this section I will argue that EM meets Carnap's criteria for successful explication. I will take the criteria of similarity, exactness, fruitfulness and simplicity in turn. For each criterion I will argue that EM meets that criterion successfully, and will also consider the extent to which alternative explications of emergence meet the same criterion.

6.1 Similarity

In Carnap's words, "*The explicatum is to be similar to the explicandum in such a way that, in most cases in which the explicandum has so far been used, the explicatum can be used; however, close similarity is not required, and considerable differences are permitted.*"¹⁶³ The question at hand, then, in asking whether or not EM is sufficiently similar to the explicandum, is firstly whether or not in most cases where the concept of emergence is used, the concept of emergence specified by EM could be used, and secondly whether or not in most cases where EM applies, the concept of emergence is used.

The answer to the first question is yes, with some unproblematic exceptions. The most serious exception is those metaphysical accounts of emergence that portray emergents

¹⁶³ Carnap, R. (1950) pg 7

as necessarily metaphysically autonomous from their micro-level bases. EM doesn't rule out a metaphysical interpretation of certain cases of emergence, however, but presents a conception of emergence according to which not all cases of emergence obtain for metaphysical reasons. According to EM, emergence itself is not a metaphysical concept, but some cases of emergence may obtain for metaphysical reasons. Another version of this objection is the claim that relativization to an observer rules out those treatments of emergence that portray emergence as obtaining for all observers (which includes metaphysical accounts of emergence). According to such treatments, if any property is emergent then it is emergent for all. In response I would point out that EM doesn't rule out cases in which some macro-level property emerges from a given set of micro-level properties for a great many observers, or even all observers, and so such cases don't raise problems for EM.

Someone wedded to a metaphysical concept of emergence may, however, take further exception to EM along the following lines: common practice shows that the only cases of emergence worthy of the name are those in which the property is emergent for all observers, for metaphysical reasons. The proponent of this objection would argue that, given this feature of the concept of emergence, the problem with EM is that it permits weak cases of emergence that aren't interesting or strong enough to deserve the name "emergence". In response I would argue that my research into common use of the concept of emergence in philosophy and science shows that not all, or even most, uses of the concept presuppose that the emergence in question must be strong. Instead, this research shows that there are two dominant conceptions of emergence, epistemic and metaphysical. These share some features

– connection to explanation, connection to failure of deducibility or derivability – but not the idea that emergent autonomy should be understood entirely in metaphysical terms.

Furthermore, I have offered a reconstruction of the concept of emergence as relative to an observer, but this does not rule out the possibility that some cases of emergence may obtain for metaphysical reasons. In such cases, the relevant property would simply be emergent for all observers. I leave open the option of giving a metaphysical interpretation of such cases – my only restriction is to demand that we understand emergence as relative to a perspective. This rules out taking emergence *itself* as a perspective-dependent metaphysical relation, but it doesn't preclude developing a metaphysical interpretation of particularly intriguing, prevalent cases of emergence if there is good reason to do so.

Finally, let's test two alternative explications of emergence against Carnap's similarity criterion. I will examine a metaphysical explication of emergence, according to which a macro-level property is emergent only if it is metaphysically autonomous from the micro-level properties, and an epistemic explication of emergence, according to which emergence is a merely epistemic phenomenon. To make these alternatives as charitable as possible, I keep them quite close to EM, tweaked only to encompass the relevant epistemic and metaphysical differences. The alternative explications are as follows:

(MET) Given components A, B, C...n arranged in relation r into a whole, a property x of the whole is emergent iff

- i) there is no available scientific explanation of the fact that the following regularity obtains of natural necessity: *Whenever components A, B, C...n are combined in relation r, the resulting whole instantiates property x*

- ii) because x has causal powers not had by any property of A, B, C...n when not arranged in relation r.

(EP) Given components A, B, C...n arranged in relation r into a whole, a property x of the whole is emergent iff

- i) there is no currently available scientific explanation of the fact that the following regularity obtains of natural necessity: *Whenever components A, B, C...n are combined in relation r, the resulting whole instantiates property x*
- ii) though an explanation will become available as our knowledge improves.

I include the notion of causal powers in MET because it is a standard way to understand metaphysical autonomy, but one could replace this clause with a different metaphysical notion of autonomy, such as failure of reduction or modal differences. The important idea is that MET portrays emergent autonomy in metaphysical terms. The epistemic criteria are included in EP to indicate that emergence exists in virtue of a contingent gap in knowledge, which is a feature of most epistemic accounts of emergence.

I have already argued that an explication like MET doesn't successfully meet the similarity criterion because it cannot be used in contexts where emergence is taken to be a contingent epistemic phenomenon, and given that such contexts include many scientific uses of the concept, this makes MET particularly unscientific as well as insufficiently similar. Alternatively, EP also fails to meet the similarity criterion because it cannot be used in contexts where emergence is turns out to be a permanent feature for all observers. EM avoids this problem because relativization to an observer makes room for the convergence of many observers on many different cases of emergence, whereas EP doesn't include this element

and furthermore a core element of EP is the idea that the unavailability of explanation characteristic of emergence is contingent. EM fares better than both of these alternative explications against the similarity criterion.

6.2 Exactness

The criterion of exactness demands that, as Carnap puts it, “*The characterization of the explicatum, that is, the rules of its use... is to be given in an exact form, so as to introduce the explicatum into a well-connected system of scientific concepts.*”¹⁶⁴ In our case, we wish to introduce the explicatum into a well-connected system of *philosophical* concepts, as well as scientific. An appropriate test, then, for how well EM meets the criterion of exactness in comparison with its rivals, is to ask if in most cases it would be clear whether or not the concept applies, and if the explicatum is capable of being introduced into contemporary science and philosophy and thereby connecting with theories in those disciplines.

EM involves the unavailability (the meaning of which is specified by the observer) of a scientific explanation (also specified by the observer). Accordingly, to find out whether some property is emergent, we must find out what the relevant observer is looking for and what they take “unavailable” to mean. This is fairly simple, and given that the explanation condition is shared across EM, MET and EP, if it is inexact in one of these reconstructions it will be so across all. The concept of “explanation” is common in science and philosophy, and so would pose no barrier to EM connecting with theories across those disciplines, and the same is true for the concept of “available”. The part-whole relation depends on the details of the case at hand, but as I mentioned before is to be interpreted very loosely and so requires

¹⁶⁴ Carnap, R. (1950) pg 7

no new conceptual apparatus to understand. Overall, it would be easy in any particular case, given knowledge of the relevant facts, to work out whether some property is emergent according to EM, and none of the concepts invoked in EM pose a barrier to connections with scientific or philosophical theories.

Let's now assess MET and EP against the exactness criterion. Note that, because MET includes a difference in causal powers as a criterion for emergence, the proponent of MET will face significant epistemic problems when applying this concept. As the case of the British Emergentists famously shows, it can be very easy to attribute a robust metaphysical difference to the world that turns out to be merely epistemic on the discovery of new information about the phenomenon in question.¹⁶⁵ Given this feature, it isn't clear in any given case whether or not MET applies, and so MET fares poorly on the exactness criterion. The causal condition may also raise problems when it comes to connecting with scientific theories about emergence, which often do not invoke metaphysical notions such as causation.

EP will fare about as well as EM on the exactness criterion, as it too involves the unavailability of an explanation. Overall, then, it looks as though EP and EM meet the exactness criterion fairly well, and that MET does not.

6.3 Fruitfulness

Carnap says of the fruitfulness criterion "*The explicatum must be a fruitful concept, that is, useful for the formulation of many universal statements...*"¹⁶⁶ Someone familiar with the philosophical literature on emergence may argue that my explication fails to meet this

¹⁶⁵ McLaughlin, B. (1992)

¹⁶⁶ Carnap, R. (1950) pg 7

criterion, for the following reason: this explication fails to make room for substantial philosophical debate about emergence. If we accept this explication of emergence, then we are committed to the idea that all we need to do to find out if some macro-level property is emergent is to ask whether or not there is an explanation of a certain sort available to some observer. Further philosophical debate is halted in its tracks, and so this is not a fruitful explication of emergence.

In response, I acknowledge that it is an implication of EM that the question of whether some property is emergent or not for some observer is simple to answer. However, I see this as a positive feature. In making the question of whether or not some phenomenon is emergent so easy to answer, this explication rules out non-substantial debates about emergence in which people starting with different definitions of emergence talk past each other. EM also opens up avenues for fruitful research into the nature of emergence, continuous across different areas in science and philosophy, because EM offers such a versatile, unified conception of emergence. For example, philosophers can address themselves to questions about exactly why a given macro-level property is emergent for some observer, whether this is true for other observers and for what reasons. They can also address themselves to questions about scientific explanation; much of the debate about the relationship between properties of the brain and properties of qualitative experience has been occupied with discussion of exactly which sort of explanation we ought to take as appropriate for this relationship.¹⁶⁷ This research is continuous with and complimentary to empirical research into the reasons for different cases of emergence, which can be conducted by

¹⁶⁷ See the Chalmers, D. (1996) and Block, N. & Stalnaker, R. (1999)

scientific practitioners. This level of unification and continuity across different approaches to emergence is unavailable to the proponent of MET or EP.

EM portrays emergence as prevalent and often cheap, but it allows for continuous dialogue and research into emergent phenomena on the part of philosophers and scientists. EM provides a framework within which reasons for the presence of emergent phenomena can be researched and understood, which meets Carnap's criterion of "*useful for the formulation of many universal statements*".

MET fares badly on the fruitfulness criterion. In endorsing a metaphysical explication of emergence we effectively embrace the idea that further empirical research won't add to our philosophical understanding of the nature of the emergent phenomenon. Once some phenomenon has been labeled as metaphysically emergent, empirical research effectively stops. Indeed, in this particular case, metaphysical research into the reason for a given case of emergence also stops, given MET's causal powers requirement. Furthermore, because MET requires that all emergence be metaphysical, most of the cases that are labeled as emergent in scientific practice will no longer count as emergent, and the potential for dialogue between philosophers and scientists that is supported by EM is unavailable to the proponent of MET.

EP fares slightly better on this criterion than MET. On an epistemic conception of emergence, we assume that the macro-level property is emergent for *merely* epistemic reasons, rather than metaphysical reasons. This allows for and even encourages further empirical research into the natural phenomena involved in purported cases of emergence. It does not, however, allow for philosophical research into the possibility of emergence obtaining for metaphysical reasons, and so EP faces a reversed version of the problem that

faced MET. MET made little room for scientific research into emergence, and EP makes little room for philosophical research. By permitting different cases of emergence to obtain for very different reasons, and by allowing for continuous research and dialogue about emergence across philosophy and the sciences, EM meets the fruitfulness criterion more successfully than EP or MET.

6.4 Simplicity

The simplicity criterion states that the explicatum must be as simple as meeting the other three requirements will allow. Someone may argue that EM is not simple, perhaps pointing to the relativization to an observer as a rather complicated aspect of EM, but given how central that aspect of the reconstruction is for enabling the explicatum to meet other criteria, fruitfulness and similarity in particular, this consideration doesn't hold much weight.

MET fares worse by the simplicity criterion than EM because the causal aspect of that explication makes it more complicated. EP fares roughly the same as EM on this criterion; the contingent element makes EP slightly more complicated than EM but the lack of relativization to an observer in EP balances out this consideration.

Overall, a survey of Carnap's criteria indicates that the rational reconstruction EM meets these criteria, and does so more successfully than other mainstream explications of emergence. EM fared better than the MET on all criteria, better than EP on two criteria and equally well as EP on two criteria. In Section 7 I will consider and respond to some objections to EM as an explication of emergence.

Section 7: Objections

7.1 Neither metaphysical nor mysterious

Someone could argue against EM along the following lines: the philosophical community may not agree on much about emergence, but it agrees on the idea that emergence is mystical and surprising. EM portrays emergence as a prevalent, cheap phenomenon and so fails to do justice to mainstream conceptions of emergence. Alexander is famously quoted by Broad as describing emergence as a phenomenon that we must accept unexplained, “*with the philosophical jam which Professor Alexander calls ‘natural piety’*”¹⁶⁸, and yet my explication of emergence captures none of this mystery – it requires no piety.

In response, I would argue first of all that the conception of emergence as inherently mysterious is not universally shared across either the philosophical or scientific literature on the subject. Epistemic and scientific conceptions of emergence, for example, typically don’t involve this appeal to mystery, though they may portray emergence as often interesting and unpredictable. Even if this feature were a universal commitment, however, this consideration doesn’t make for an objection to EM as an explication of emergence because EM allows for a certain amount of mystery. Obviously not all cases of emergence, according to EM, will be mysterious. But those cases that are prevalent across many different observers and are also stable for those observers, and in which the reasons for the emergence are poorly understood, will remain both intriguing and mysterious.

¹⁶⁸ Broad, C.D. (1925) pg 55.

Furthermore, one could argue that in EM the sense of mystery associated with emergence is replaced with the idea that emergence involves the unavailability of a certain kind of explanation. A plausible way to think about this replacement is that when we originally thought emergence was mysterious what was actually involved was a failure of an attempt at an explanation.

One kind of mystery that EM makes no room for is the kind associated with the idea that emergence is basic, brute and unexplainable-in-principle. I would argue that this is no great conceptual loss, however, as endorsing the existence of such mysterious phenomena discourages philosophical and scientific research. It is a benefit of EM that the possibility is always open that some new piece of knowledge or theoretical tool may come along and render a property that was emergent for many observers no longer emergent. An illustrative way to think of this feature of EM is in terms of the attitude it recommends that we adopt towards cases of emergence. EP, among other epistemic views on emergence, demands optimism about emergence, the idea being that what is unexplainable now will not be so in the future. MET demands pessimism, because emergence involves features that are unexplainable in principle, for metaphysical reasons. EM, on the other hand, demands neither optimism nor pessimism about emergence. Some cases may turn out to be intractable, while others may disappear quite quickly, and EM does not require us to adopt the same attitude towards all cases.

7.2 Unified, but sparse and cheap

EM definitely unifies diverse conceptions of emergence, but perhaps at the cost of content; one could argue that the only way such impressive unification can be achieved is by

making the explication almost content-free. EM is so sparse that it is hardly surprising that it can encompass almost every account of and purported case of emergence. This is a legitimate concern, and a concept of emergence that is so broad as to say almost nothing about the phenomenon is of little interest and little use.

In response, I would argue that EM does portray emergence as prevalent and often cheap, but also allows for differences between more and less interesting cases of emergence, depending on the nature of and number of observers in question, the kind of explanation they have in mind, and the relevant notion of availability. Emergence may be prevalent, according to EM, but not all cases of emergence are equal. There are many different kinds of cases of emergence, and some of them are far more interesting than others. We can bring philosophical and scientific resources to bear on questions about why a given property is emergent for a given observer or observers, and all the while practitioners can avoid the problem of “talking past each other” about emergence.

EM provides resources to differentiate between cases, and for conducting ongoing dialogue and research into the nature of emergence. This should ameliorate concerns about the cheapness and prevalence of emergence according to EM.

Section 8: Putting the explication to work

Having formulated and defended EM as an explication of emergence, I will now examine two case studies in which EM offers a clear, helpful approach to a long-standing philosophical or scientific debate. The cases are: the explanatory gap in contemporary philosophy of mind and the mechanist-vitalist debate in the history of philosophy. Examining

these cases will enable me to demonstrate that this explication of emergence can perform useful work, by cutting through confusion and clarifying debates.

8.1 The explanatory gap

The “explanatory gap” is a label used in philosophy of mind for the idea that we lack a satisfactory explanation for the fact that conscious experience arises from brain activity.¹⁶⁹

Levine offers an intuitive statement of the idea of an explanatory gap:

While we seem to have some idea how physical objects, or systems, obeying physical laws, could instantiate rational and intentional properties, we have no idea, I contend, how a physical object could constitute a subject of experience, enjoying, not merely instantiating, states with all sorts of qualitative character... There seems to be no discernible connection between the physical description and the mental one, and thus no explanation of the latter in terms of the former.¹⁷⁰

There are many different formulations of the explanatory gap. Some think that the gap is between conscious experience and physical properties, while others take the relevant gap to be between conscious experience and neuroscientific properties. Some deny the existence of such a gap, while others strongly affirm its existence. There are also differences between different authors in their views on the significance of the gap, what sort of explanation is thought to be missing, and what it means for an explanation to be missing.

The explanatory gap plays a major role in philosophy of mind because many people not only think that there is such a gap, but that its existence generates a serious problem for physicalist theories of mind. The basic idea behind this challenge is that, if physicalism is true, then a physicalistically acceptable explanation of conscious experience should be

¹⁶⁹ Although the term “explanatory gap” is most often used in philosophy of mind, the idea crops up in some other areas of philosophy, including debates about reductionism and non-reductionism in metaphysics.

¹⁷⁰ Levine, J. (2001) pg 76

available. This explanation should enable us to cross the explanatory divide between physical properties and consciousness without positing any non-naturalistic entities, properties or substances. If such an explanation is not available, then, according to this line of thought, we have good reason to think that physicalism is not true. Accordingly, the explanatory gap has become one of the major obstacles to a physicalist theory of mind, and has generated a wide range of literature in which naturalistically inclined philosophers of mind attempt to grapple with this problem.

Responses to the challenge to physicalism vary. Some argue that there is such a gap, and that it is a problem for physicalism.¹⁷¹ Some argue that there is such a gap, but that it is not a problem for physicalism.¹⁷² Others argue that the gap doesn't pose a problem for physicalism, because we can explain the gap through appeal to the idea that phenomenal experience requires particular conceptual resources.¹⁷³ And finally, some have taken the explanatory gap as a reason to endorse non-physicalist theories of mind, and to devote their attentions to making such theories as naturalistic as possible.¹⁷⁴

In many of these cases, differences between responses to the explanatory gap come about as a result of very different views on the appropriate kind of explanation to look for, and about what it is for an explanation to be "missing". It is therefore tempting, as in the case of emergence in general, to take much of the debate over the explanatory gap as involving

¹⁷¹ Jackson, F. (1982), Nagel, T. (1974)

¹⁷² Block, N. & Stalnaker, R. (1999). They argue that the explanatory gap isn't a challenge for physicalism because physicalism needn't be construed in terms of a priori analysis.

¹⁷³ This is known as the "*Phenomenal Concepts Strategy*". See e.g. Balog, K. (forthcoming)

¹⁷⁴ Chalmers, D. (1996)

philosophers talking past each other. Armed with EM, however, we can distinguish the deep, interesting debates about the explanatory gap from the insubstantial, misguided debates.

The first step in this process is to understand claims about the explanatory gap as claims about emergence, in line with the explication of emergence given in EM. On this picture, the view that there is an explanatory gap amounts to the view that for some observer/s and some property p of conscious experience, there is no explanation available to that observer/s of the fact that the following regularity obtains of natural necessity: *Whenever components N_1, N_2, \dots, N_n are combined in relation r , the resulting whole instantiates property p* , where components $N_1 \dots N_n$ can be understood as neurophysiological components.¹⁷⁵

This picture makes it clear that the following question is underspecified: “*Is there really an explanatory gap?*” If we think of the explanatory gap in terms of EM, we can see that to ask this question has suppressed subscripts that need to be given values before we can ask philosophically interesting questions, and that to ask this question without filling out the subscripts is to miss the point about the explanatory gap. We must first of all begin by filling in the subscripts, by asking, for example: “*For whom does this case of emergence obtain?*” “*What is the relevant form of explanation?*” “*What is the relevant standard of availability?*” “*Is there a case of emergence if we switch from neurophysiological components to physical components?*” These questions are not particularly deep, but need to be answered in order to fully understand the case. Having answered such questions, we can then move on to address the philosophical issues. For example: “*What sort of explanation of such a phenomenon should we be seeking?*” “*What is an appropriate standard for the availability of an*

¹⁷⁵ They could be physical, chemical etc components instead but I’ll use neurophysiological components as a case study.

explanation?” “What are the implications of there being a case of emergence for x many observers with y conception of explanation?”

In offering this analysis of the debate, I don't mean to suggest that the debates on the explanatory gap are for the most part misguided. Indeed, some of the most notable work on the subject, including the dialogue between Jackson & Chalmers and Block & Stalnaker, can be understood as addressing the following question: what sort of explanation ought we to be looking for in this case, and should that form of explanation involve a priori derivation?¹⁷⁶ However, on first exposure to the debate about the explanatory gap, it can be tempting to go straight for questions that on closer examination turn out to be misguided, such as, “*is there really an explanatory gap?*” Understanding the debate as a debate about emergence, conceived of in terms of EM, makes it clear that we should avoid such questions, and why we should avoid them.

8.2 The vitalist-mechanist debate

The debate between vitalists and mechanists is hard to pinpoint, both chronologically and philosophically. For the sake of convenience I will stipulate that I am interested in the 19th century incarnation of the debate, and will offer a rough characterization of the philosophical positions on both sides. I don't intend to offer a definitive historical interpretation of this debate, but instead wish to offer one interpretation of it as a case to which we can apply EM. This application will, as in the case of the explanatory gap, enable us to distinguish interesting, well-formed questions from uninteresting ones, and will also

¹⁷⁶ See Chalmers, D. & Jackson, F. (2002) and Block, N. & Stalnaker, R. (1999)

provide a clear framework within which to understand how the positions changed as the relevant science progressed.

The distinction between the mechanist and the vitalist can be drawn in terms of their respective approaches to the question of life. Beckermann offers a rough characterization of the distinction, which I will adopt:

“... mechanists claimed that the properties characteristic of living organisms (metabolism, perception, goal-directed behavior, procreation, morphogenesis) could be explained mechanistically, in the way the behavior of a clock can be explained by the properties and the arrangement of its cogs, springs and weights. ... vitalists, on the other hand, maintained that the explanation envisaged by the mechanists was impossible and that one had to postulate a special nonphysical substance in order to explain life – an entelechy or élan vital.”¹⁷⁷

There were different varieties of vitalism and of mechanism, depending on factors such as the nature of the mechanistic explanation in hand and the strength of the associated metaphysical thesis. Indeed, some have argued that it would be better to think of mechanism and vitalism as collections of strands of thought rather than unified philosophical positions. Benton, for instance, argues that *“it is a mistake to regard vitalism as one doctrine, or set of doctrines.”¹⁷⁸* For the purposes of this discussion, however, I will continue to use Beckermann’s definition.

C.D. Broad originally presented his account of emergentism as providing middle ground between mechanism and vitalism.¹⁷⁹ This may make it confusing to apply EM to the mechanist-vitalist debate, because doing so requires us to think of the debate in terms of emergence. But I will suggest that in this section we put aside the history of the term

¹⁷⁷ Beckermann, A. (2000) pg 1

¹⁷⁸ Benton, E. (1974) pg 17

¹⁷⁹ Broad, C.D. (1925) pg 61

“emergence” for the moment and wholeheartedly adopt EM as the correct conception of emergence. Then we can, as in the case of the explanatory gap, apply EM to the mechanist-vitalist debate.

The first step in this process, as in the case of the explanatory gap, is to conceive of vitalism as a claim about emergence. According to this view, vitalism amounts to the position that for some feature *l* (a property characteristic of living creatures) and some observer/s, there is no explanation available to the observer/s of the fact that the following regularity obtains of natural necessity: *Whenever components C_1, C_2, \dots, C_n are combined in relation r , the resulting whole instantiates property l .* Components $C_1 \dots C_n$ can be understood as chemical components, but depending on the relevant kind of mechanism, they may alternatively be understood as physical or physiological components.

To get to some of the interesting questions about vitalism, we can fill in the details. Again, we can see that the question “*Is vitalism true?*” is underspecified. There are suppressed subscripts, which must be filled out by asking questions such as “*What is the relevant notion of explanation?*” and “*For whom does this case of emergence obtain?*” Having done so, we can then begin to ask philosophically and scientifically interesting questions about this debate, including “*Were scientists right to look for mechanistic explanations of these natural phenomena, rather than an alternative sort of explanation?*” or “*What is the significance, metaphysical or otherwise, if a mechanistic explanation is unavailable for y number of observers?*”

One of the scientifically interesting questions in this area is the question of whether or not a mechanistic explanation of the relevant feature was available, and this debate actually

took place over the particular case of fermentation. Louis Pasteur and Marcelin Berthelot offered rival accounts of fermentation, with Pasteur arguing that fermentation was a vital process and Berthelot arguing that fermentation could be explained mechanistically, without needing to call upon the notion of life. Fruton quotes Berthelot stating his objective as follows: “*To banish life from all explanations relative to organic chemistry, that is the aim of our studies*”¹⁸⁰ and Berthelot claimed to have developed an entirely chemical (in our sense, mechanistic) explanation of fermentation. It is important to reconstruct this exchange as a genuine debate, rather than two scientists talking past each other, and EM provides resources for doing so. Pasteur and Berthelot shared a notion of explanation and of availability, and had a substantial, genuine disagreement about whether or not an explanation of a specific sort was available according to their shared standards.

In contemporary philosophy and science debates about the nature of life have changed, with the focus of attention switching from formulating an explanation of the instantiation of features commonly associated with life to other questions about the nature of life.¹⁸¹ We can use EM to represent this state of affairs by stating that either (depending on the motivation for this historical claim, given that there are different versions in the literature) we have uncovered a mechanistic explanation, thereby making an explanation available and properties such as I no longer emergent, *or* we uncovered some other form of explanation and decided not to look for mechanistic explanations of features associated with life any more. Alternatively, some may hold that this shift in attention hasn’t resulted from any developments in the mechanist-vitalist debate but is simply a product of different trends

¹⁸⁰ Fruton, J (2006) pg 61

¹⁸¹ Including attempts to develop improved definitions of life, and research into the limits of life involving A-Life.

taking hold in philosophy and science, in which case there is no change in the debate to represent. Either way, looking at this debate through the lens of EM provides clear distinctions between interesting and uninteresting issues, preserving the idea that disagreements between scientists of the time were substantive and providing a simple way to understand the historical development of the mechanist-vitalist debate.

Section 9: Conclusion

Following a brief survey of philosophical and scientific approaches to emergence, I offered an explication of the concept of emergence as the unavailability of a scientific explanation for an observer or observers. I argued that this explication meets Carnap's criteria for successful explication and that it has substantial unificatory appeal, as it allows for continuity between philosophical and scientific research into emergence.

Endorsing this view requires a shift in the way we think about emergence. The interesting questions about emergence are not, on this view, questions about whether some property is emergent, but instead are about who the property is emergent for, what their conception of scientific explanation is and so on. Instead of regarding emergence as a homogenous phenomenon, as it is typically understood, my explication of emergence portrays emergence as prevalent and very diverse. This change may be radical, but in making

it we replace a “*notorious philosophical term of art*”¹⁸² that “*often causes confusion in science and philosophy*”¹⁸³ with a precise, unified and naturalistic concept.

Having offered this explication of the concept of emergence, I will now move on to formulate and defend my positive account, the *perspectival account of emergence*, in Chapter 5.

¹⁸² O’Connor, T. (2006)

¹⁸³ Chalmers, D. (2006) pg 1

Chapter 5: *The Perspectival Account of Emergence*

Section 1: Introduction

In this chapter I will formulate and defend an account of emergence, the *perspectival account of emergence*, based on the results of Chapters 2 through 4. In Section 2 I will formulate the account and discuss some of its central features in detail. In Section 3 I will explore the metaphysical implications of the perspectival account of emergence. In Section 4 I will describe a range of different ways to differentiate between the many cases of emergence permitted by the perspectival account of emergence. In Section 5 I will consider the problem of downward causation, which is traditionally thought to pose problems for accounts of emergence. And finally in Section 6 I will consider some objections and replies to the perspectival account of emergence.

Section 2: The Positive View

In Chapter 4, I introduced a distinction between the tasks of formulating a rational reconstruction of the *concept* of emergence, and formulating a positive *account* of emergence. For a reconstruction of the concept of emergence to be successful, there need be no actual cases of emergence in the world, but merely people using the concept, and

standards for a good reconstruction¹⁸⁴ are based on factors such as accuracy to common use of concept and potential benefits associated with adopting the reconstruction. A positive account of emergence, on the other hand, not only picks out the correct *concept* of emergence, but must also address questions about whether there *is* anything that corresponds to that concept, and if so, where that it stands in the metaphysical landscape. A good example of this distinction comes up in the case of causation. A reconstruction of the concept of causation may include features such as time-asymmetry or transitivity, while a positive account of causation must say something about whether such features are parts of the metaphysical landscape, or whether they are epistemic features, and so on.

In this chapter I will offer a positive account of emergence based on the rational reconstruction of the concept of emergence developed in Chapter 4, which I called “EM”. There are a great many cases of the phenomenon picked out by EM and so there is no hard work to be done in arguing that this emergence *exists*. There remain, however, hard questions to answer about the metaphysical implications of this account of emergence and about how to differentiate between different cases of emergence, as well as a number of objections to consider, and Sections 3 to 6 of this chapter will be devoted to responding to those questions and objections.

In Chapter 4, I offered the following rational reconstruction of the concept of emergence:

(EM) Given components A, B, C...n arranged in relation r into a whole, and an observer O, property x of the whole is emergent for O iff there is no scientific

¹⁸⁴Such as those developed by Carnap in Carnap, R. (1950)

explanation available to O of the fact that the following regularity obtains of natural necessity: *Whenever components A, B, C...n are combined in relation r, the resulting whole instantiates property x.*

My positive account of emergence, the perspectival account of emergence, is as follows:

A property is emergent iff it meets the conditions of (EM).

EM portrays emergence as relative to a number of different factors: an observer, a component-whole relation, occupants of that component-whole relation, a kind of explanation, and a notion of unavailability. I will now briefly describe each of these factors.

I understand “observer” to simply mean any live human observer. This rules out cases of emergence that are relative to the point of view of, for example, Laplacean Martians or computer programs.¹⁸⁵ I don’t impose any restrictions on the notion of observer other than that observers be live humans, and this view accordingly permits cases of emergence relative to the points of view of children, people who cannot perform mathematical calculations and so on. I will address the challenges raised by such cases in Section 6 of this chapter, and for the moment I will simply point out that I do permit such cases as cases of emergence. I endorse the claim that many of the cases in which emergence is relative to an observer in a particularly unusual or epistemically disadvantaged situation make for deviant or uninteresting cases of emergence, but later on I will argue that I have resources to accommodate this fact without imposing further restrictions on what counts as an observer.

¹⁸⁵ I find this to be the most sensible way of picking out observers, but there is nothing to prevent someone from tweaking my account of emergence to include a somewhat looser notion of “observer” that would permit such cases.

The second and third factors to which emergence is relative, according to EM, are a component-whole relation and the occupants of that relation. I spent some time discussing the different contenders for component-whole (or micro-macro) relation in Section 2 of Chapter 1, so for the moment I will simply point out that I take an ecumenical position on the nature of the component-whole relation, and leave it open with the aim of encompassing as many current uses of the concept of emergence as possible. In doing so I run the risk of permitting deviant cases of emergence, as before, but I will argue that I have resources on hand to deal with this problem. To specify the *occupants* of the component-whole relation is simply to pick out the phenomenon of interest, whether it is the relation between properties of a particular chemical compound and properties of its component elements, or the relation between properties of cells in an Artificial Life configuration and properties of the configuration, or some other relation between micro- and macro- levels.

The fourth factor to which emergence is relative, according to EM, is a kind of explanation. Rather than specifying a particular notion of explanation to which emergence is relative, In Chapter 3 I argued that it is best to leave this open. If I specify one kind of explanation to which emergence is relative, then the appeal to explanation that I introduced in Chapter 3 will not offer a general solution to the collapse problem, because emergence is associated with the unavailability of many different kinds of explanation across the philosophical and scientific literature. See Sections 2 and 5 of Chapter 3 for a further discussion of this claim.

The final factor to which emergence is relative, according to EM, is a notion of unavailability. This is included to capture the idea that different investigators will work with different notions of unavailability depending on disciplinary norms, among other factors. For

example, a group of epistemologists may hold a standard of unavailability such that an explanation is only unavailable if an ideal epistemic agent couldn't formulate one. On the other hand, a group of neuroscientists may hold that an explanation is unavailable if the best neuroscientists in the world have been attempting to formulate one for a decade or more and have not yet succeeded. I have placed no restrictions on what counts as unavailability, which as before generates a "deviant cases" problem that I will address later on in this chapter.

In Section 3 I will examine the metaphysical implications of the perspectival account of emergence, but before doing so I will offer some general reflections on the difference between this perspectival account of emergence and a metaphysical account of emergence. The perspectival account of emergence portrays emergence as relative to an observer, and hence not a part of the mind-independent metaphysical landscape. Note, however, that a standard way to move from such a relativistic account to a metaphysical account would be to argue that there is a metaphysically privileged version of each one of the relevant factors (explanation, unavailability, observer and so on), and that cases of emergence relative to *those* factors are the real, metaphysical cases of emergence. I will discuss this move in some detail in the following section, but for the moment I simply wish to point out that I do not intend to make it. I will show in Section 3 that the perspectival account of emergence is compatible with, though does not entail, a variety of different metaphysical views.

In this section I have discussed each of the factors to which emergence is relative: an observer, a component-whole relation, occupants of that component-whole relation, a kind of explanation and a conception of unavailability. Each of these factors also permits a different way of individuating and classifying different cases of emergence, and in Section 4 I will explore some other ways to individuate between different cases of emergence.

Section 3: Metaphysical Implications

The perspectival account of emergence portrays emergence as relative to a number of different factors, most notably to an observer. In this section I will expand on this relativism, and discuss its implications for the metaphysical status of emergence.

3.1 Observers and Perspectives

The definition of an “observer” in account is simply a live human being, and I don’t impose any further restrictions on what it takes to be an observer. “Observers”, for the purposes of this account of emergence, are individuated along the same lines as live human beings. This simple mode of individuation enables the perspectival account of emergence to avoid difficult questions about what “perspectives” are, how best to individuate perspectives, and what it takes to occupy one.

There is a broader notion of “perspective” which may appear to be a more attractive resource for the development of an account of emergence than this simple idea of an “observer”. I won’t attempt to offer any kind of analysis of the concept of a perspective, but even without doing so it is clear that there are some differences between the general idea of a perspective and the idea of an observer, understood as a live human being. For example, a perspective can be shared between multiple observers, such that we can speak of the “perspective of women in philosophy”, or the “perspective of 19th century chemists”. People can also occupy more than one perspective at once; for instance, I occupy the perspective of a denizen of Chapel Hill and the perspective of a citizen of the United Kingdom. This general, rough notion of “perspective” doesn’t come attached to a simple mode of individuation, and it isn’t clear what differentiates one perspective from another.

Although it is simpler to base an account of emergence around observers, rather than perspectives, because it is simpler to individuate observers and also to say what they are, it is important to be able to capture the idea that in some cases, a property is emergent for multiple observers. An account of emergence as relative to a perspective rather than an observer would allow for this feature because multiple observers can share perspectives. Rather than appealing to perspectives, however, I capture this feature through appeal to the idea that multiple observers may *converge* on particular cases of emergence. Some cases of emergence may be emergent for only one observer, while others may be emergent for a great many, or even all, observers. This convergence feature allows us to capture the sensible idea that sometimes, particularly in situations where observers work together in groups, certain properties will be emergent for the group, and not just for some particular individual. The appeal to convergence allows us to accommodate such cases without having to grapple with the abstract, vague idea of perspective.

In line with conventions in the literature, however, I will continue to call this account a “perspectival” account of emergence rather than, say, an “observer-relative” account of emergence. This is because the term “perspectival” is often used to describe views that take some entity or relation to be a feature of interactions between people and the world, rather than a strictly mind-independent feature of the world, and this is the category in which the perspectival account of emergence belongs.

The appeal to observers generates a simple question: When an observer gains new knowledge (or a tool etc) so that a property is no longer emergent for them, does that mean that the *observer* relative to which this property was emergent has now changed? Or is it instead the case that the observer is the same, but that the property is no longer emergent for

that observer? I opt for the latter answer. If some development occurs such that a property that was emergent for a particular observer is no longer emergent for that observer, then that property *was* in the past emergent for that observer but that it is now no longer emergent for that observer. This way of representing such cases enables me to avoid counterintuitive claims such as that the *identity* of the observer changes in light of having learnt something new.

3.2 *Metaphysical Implications*

Embracing the perspectival account of emergence involves understanding emergence as relative to an observer. According to some alternative, metaphysical accounts of emergence, emergence is a relation that obtains between groups of properties that are metaphysically distinct. The idea behind this picture of emergence is that emergence tracks distinctions in the mind-independent universe, and understanding emergence as relative to an observer interferes with this conception of emergence. In this section I will expand on the original position and on the nature of this interference.

Mainstream metaphysics is a common term for a conception of metaphysics according to which the goal of metaphysics is to delineate the structure of the mind-independent universe.¹⁸⁶ Most of those who embrace this conception of metaphysics believe that such enquiry is possible, and hence that we can limn the structure of the mind-independent universe. On the mainstream conception of metaphysics, anything that isn't part of the mind-independent universe isn't a proper subject of metaphysical enquiry, though it may be a proper subject of philosophical enquiry. To think of emergence as metaphysical in

¹⁸⁶For a thorough illustration of this position, see Manley, D. (2009) and Sider, T. (forthcoming)

this sense is to think of emergence as a relation that obtains in the mind-independent universe.

Embracing the idea of emergence as relative to an observer involves embracing the idea that emergence is not part of the mind-independent universe, and so not properly “metaphysical” in mainstream terms. As I have presented the perspectival account of emergence, there are no facts of the matter about whether or not there is any emergence independently of observers. Of course, whether or not some property is emergent from some other properties *for* a certain observer may be a perfectly objective matter, but without the observer there is no fact of the matter about emergence.

Of course, many philosophers are quite happy with the idea that metaphysical enquiry may include enquiry into features of the universe that appear to be dependent upon the interests of human observers, on what they find interesting and choose to explain. The definition of metaphysics, and of the proper role of metaphysical enquiry, raises questions that beyond the scope of the present discussion. In the mean time I will simply say that, according to one popular, contemporary conception of metaphysics, understanding emergence as relative to a perspective puts emergence beyond the domain of metaphysical enquiry, because it portrays emergence as relative to an observer, rather than as mind-independent.

The position that emergence is relative to an observer also has significant implications for the causal status of emergence, which I touched on at the end of Chapter 3 but will now turn to in more detail. According to proponents of metaphysical accounts of emergence, much of what is interesting about emergence is tied up in the idea that emergence

may somehow enable us to develop a naturalistic form of metaphysical non-reductionism. On the one hand, some have argued, reductionism is unrealistic because it fails to vindicate the appearance that macro-level entities, and in particular the entities of the special sciences, are real and causally efficacious.¹⁸⁷ On the other hand, most philosophers would like to hold on to the belief that such entities are also related to the micro-level entities that compose them in, as Wilson puts it “*a naturalistically respectable way*”¹⁸⁸. Emergence has been presented as the relation that can reconcile these seemingly opposing appearances, in so far as it can account for the causal efficacy of macro-level entities without having to posit non-natural entities or properties.

For an appeal to emergence to play such a role, however, emergence must be capable of endowing macro-level entities with their autonomous causal powers. According to such an account, macro-level entities would be causally autonomous *in virtue of* having causal powers. The idea that emergence is relative to an observer poses a problem for this claim. If emergence is relative to an observer, then whether or not some property is emergent can vary from observer to observer. Few metaphysical accounts of causation portray causal capacity as varying across observers in this way, so it seems unlikely that the perspectival account of emergence can play this role of accounting for the apparent causal autonomy of macro-level entities.

The implications of this finding depend, however, on the relevant view of causation. A perspectival, subjectivist or Kantian account of causation would, depending on the details of the account, probably be able to reconcile the idea of emergence as relative to a

¹⁸⁷See Fodor, J. (1974) and Wilson, J. (2010)

¹⁸⁸ Wilson, J. (forthcoming)

perspective with the idea of emergence as the locus of the causal autonomy of macro-level entities. But most metaphysical accounts of causation, which portray causation as a mind-independent relation, would not be able to reconcile the idea of emergence as relative to an observer with the idea of macro-level entities having causal powers in virtue of having emergent properties.

Overall, then, the perspectival account of emergence has two important metaphysical implications. First of all, it portrays emergence as relative to an observer, rather than mind-independent, and so not properly “metaphysical” according to at least one popular conception of metaphysics. Secondly, it renders one popular view of the proper role for an account of emergence, as a way to vindicate the seeming causal autonomy of the special sciences, untenable on most accounts of causation.

3.3 Compatible with reductionism and with non-reductionism

I have argued that emergence is not a feature of the mind-independent metaphysical landscape. It may be tempting to interpret this view as supporting metaphysical reductionism or eliminativism of some kind. In this section I will discuss the fact that the perspectival account of emergence presents emergence as a merely explanatory, rather than metaphysical, phenomenon, but remains agnostic on the broader question of the viability of non-reductionism. I will describe how both the non-reductionist and reductionist could adapt the view to suit their own needs, although the perspectival account of emergence doesn’t entail or even lend support to either position.

So far I have argued that emergence is an explanatory, rather than metaphysical relation. This means that we cannot use the concept of emergence as a way of fleshing out, or

giving substance to, metaphysical non-reductionism. However, a proponent of metaphysical non-reductionism could use the perspectival account of emergence as a *guide* to potentially metaphysically interesting cases of emergence. For example, consider a case in which some property is emergent relative to all observers, relative to some particular kind of explanation. If the kind of explanation is thought by the non-reductionist to have a particular metaphysical status (as, for example, a priori reductive explanation is thought to have by some philosophers of mind) then they may take this as reason to think that this case warrants further metaphysical research. However, in such a case the metaphysical work will *not* be done by emergence, but by some other metaphysical relation.

Identity provides a good example of such a case. Some philosophers have argued that identity is not an explanatory relation.¹⁸⁹ If this is true, then identities will present cases of emergence (a position that appears strange at first if we consider that emergentism about the mind is traditionally opposed to identity theory about the mind, although thinking of emergence in explanatory terms deals with this counter-intuitive appearance). These cases of emergence will obtain for metaphysical reasons, in so far as identity is a metaphysical relation, but the metaphysical “work” is performed by the *identity* and not by emergence itself. The emergence in this case is effectively “underpinned” by a metaphysical relation.

I have argued that emergence is not metaphysical but this does not rule out the prospect that some cases of emergence may obtain for metaphysical reasons. Accordingly, the perspectival approach to emergence is compatible with metaphysical non-reductionism, though incompatible with a view according to which emergence is the metaphysical relation

¹⁸⁹ Thanks to Karen Neander for this case. For discussion of the idea that identities are not explanatory, see Papineau, D. (2011)

that grounds non-reductionism. The non-reductionist could even use the perspectival approach to emergence as a way of identifying cases that are metaphysically interesting, though to do so they must provide an independent argument for the position that the absence of a certain kind of explanation for particular observer or observers is a privileged guide to metaphysics.

On the other hand, the metaphysical reductionist may also use the perspectival account of emergence for their own purposes, though just as in the case of the non-reductionist, they must provide some extra independent arguments for that position. The reductionist could argue that the perspectival approach to emergence helps us to identify explanatory “gaps” (in a loose sense), which are simply cases in which we lack a decent scientific explanation of some phenomenon. The reductionist could then argue that such explanatory gaps are no guide to metaphysics, just as the non-reductionist could argue that some such gaps are a guide to metaphysics.

The perspectival account of emergence is compatible with metaphysical non-reductionism and reductionism, though neither view is entailed or even implied by the perspectival account. In each case a separate argument must be made addressing the question of whether or not certain kinds of explanatory gap are a legitimate guide to metaphysics.

In passing, we should note that this is another instance of the useful house-clearing work done by the perspectival account of emergence. Having provided tools to enable us to carefully and precisely catalogue the many different cases of emergence, we have also precisified exactly what a reductionist or non-reductionism would then need to do in order to

vindicate their position: argue that the unavailability of certain explanations either is (in the non-reductionist case) or is not (in the reductionist case) a guide to metaphysics.

The perspectival account of emergence is relativistic, and portrays emergence as non-metaphysical. However, once all of the subscripts are filled out – the part whole relation, the kind of explanation, the notion of unavailability and so on – there is a fact of the matter about whether or not some property is emergent, relative to those subscripts. This means that there can be substantive disagreement about whether or not some property is emergent, so long as such disagreement takes place against shared standards for unavailability, explanation and so on. For example, consider the case of the vitalist-mechanist controversy discussed in Section 8 of Chapter 4. I argued that reconstructing the debate between Pasteur and Berthelot as a debate about emergence presents it as a substantive debate about whether or not an explanation of a particular kind was available, relative to shared standards. The perspectival account of emergence may be relativistic, but we may still have substantive debates about emergence, so long as they are relative to shared standards.

Section 4: Differentiation among emergents

The perspectival account of emergence permits a wide range of cases of emergence. For example, one property may be emergent because the relevant observer is looking for a mechanistic explanation and none is available, while another may be emergent because the observer in question is looking for an a priori explanation and none is available. Typically accounts of emergence do not permit such a wide variety of cases, but I will argue that the fact that the perspectival account of emergence admits such a range of cases is not a problem

for that account because there are resources available to differentiate between cases. The perspectival account of emergence admits a great many cases of emergence, but does not treat them all as equally interesting. In this section, I will describe some of the different ways in which we can differentiate between cases of emergence.

An initially quite tempting way to differentiate between cases of emergence is to place cases of emergence on a spectrum running from “strong” to “weak”, where placement on that spectrum would be determined by the number of observers an emergent is emergent *for*. This is, however, misguided for a number of reasons. The idea of emergence as lying on a spectrum from strong to weak masks the wide range of reasons for which certain cases of emergence obtain, independently of the number of observers that converge upon them. Accordingly, I recommend that we should take there to be a wide range of emergents which we can differentiate in terms of different factors depending on our interests, rather than a spectrum of emergents running from strong to weak.

In this section I will survey a few obvious contenders for individuating factors, or ways to differentiate between the many different emergent properties. Note, however, that I take there to be no absolutely privileged distinctions between cases of emergence. Cases of emergence are more and less interesting depending on our goals and interests. Consider, for example, a historian of philosophy who is interested in cases of emergence that are emergent for a certain historical group for scientists, but not for contemporary scientists. Now compare her with a neuroscientist who is interested in cases of emergence that obtain only when the observers are looking for mechanistic explanations. These different people have different interests, which will dictate exactly how they differentiate between cases of emergence. The

first person will differentiate on the basis of the relevant observers, while the second will differentiate in terms of the relevant explanation.

As I mentioned in the introduction to this section, an obvious contender for differentiating between different cases of emergence is in terms of *numbers*. Some cases of emergence are emergent for many observers, while others will be emergent for only a few. Following the locution introduced in Section 2, when one case is emergent for multiple observers I will say that those observers “converge” on the case of emergence. This feature does not, however, tell us very much about different cases of emergence. Consider, for example, a case of emergence on which many observers converge because it is impossible for humans to develop the relevant kind of explanation. Alternatively, consider a case of emergence on which many observers converge out of a contingent accident; it just so happens that so far no one has developed an explanation of the relevant sort. Thinking in terms of simple numbers provides us with a starting place for differentiation, but cannot reveal much about the differences between such pairs of cases.

Another differentiation factor is the *stability* of the case of emergence for the relevant observer. The perspectival account of emergence portrays emergence as relative to an observer, but some cases will stay with an observer over time and some will not. For instance, in a case of emergence in which it is impossible for a human observer to formulate an explanation will be stable for that observer, i.e. the relevant macro-level property will be emergent for that observer for all time. A case in which the observer lacks some important computational tool which is later made available to them will be less stable; when the observer is able to use the tool and so formulate an explanation, the property will no longer be emergent for them. Cases in which the observer simply makes logical mistakes will also

be unstable for that observer, because when the observer gets better at logic, the property will no longer be emergent for them.

Another way to differentiate between cases of emergence is to differentiate on the basis of the relevant notion of *availability* in question. Some cases of emergence may be emergent for an observer with a fairly lax standard of availability, such that an explanation is unavailable if the observer herself isn't aware of such an explanation. Other cases may be emergent for an observer with a different standard for availability, such that only if it is impossible to formulate an explanation is the explanation unavailable. Cases of emergence also differ in terms of the *tools* available to the relevant observers. Some properties may be emergent for a group of observers because they lack a particular tool, such as a kind of computer. Other properties may be emergent for a group of observers because they lack certain *skills*, such as mathematical or logical skills, that would permit them to formulate an explanation of the relevant sort. These cases suggest different modes of individuation between cases of emergence, in terms of the tools and knowledge available to the relevant observers.

There are many other ways to differentiate between the many cases of emergence that are permitted by the perspectival account of emergence. I have surveyed only a few different factors that we may use to taxonomize these different cases of emergence, and one could use many more, depending on one's interests.

Section 5: Downward Causation

“Downward causation” is the term for a phenomenon whereby an emergent property has causal powers not only to affect other properties or entities at the macro-level, but also to affect micro-level properties and entities. As Bedau and Humphreys put it, in cases of downward causation, “*emergent phenomena have novel effects on their own emergence base.*”¹⁹⁰ Some philosophers have argued that downward causation is both a necessary feature of and a barrier to the existence of emergence. Kim argues for the position that downward causation is necessary for emergence, and that downward causation is impossible, hence undermining the very possibility of emergence.¹⁹¹ In this section I will briefly outline the idea of downward causation and why it is often thought to be a necessary but also problematic feature of emergence, and will then show that downward causation is not a feature of emergence according to the perspectival account of emergence.

Given an idea of emergence as a combination of dependence and autonomy, some philosophers, particularly those who favor metaphysical accounts of emergence, have argued that there must be a causal element to emergent autonomy. According to such a view, any emergence worthy of the name must involve macro-level features that are in some sense causally autonomous. As Kim puts it, “... *the causal powers the emergents bring with them must be new and distinctive... if they were reducible to the causal powers of the base-level*

¹⁹⁰ Bedau, M. and Humphreys, P. (2008) pg 6

¹⁹¹ Kim, J. (2006)

properties they would be bringing nothing new and would have nothing new to contribute to the evolving causal structure of the world."¹⁹²

If emergent macro-level properties are causally autonomous, however, in so far as their causal capacity is not exhausted by the causal capacity of underlying micro-level properties, then we face an apparent overdetermination problem. To illustrate this problem, consider two sets of properties, P and Q, where P-properties are micro-level properties and Q-properties are macro-level properties, and imagine that all Q-properties are emergent from P-properties. Q^1 is an emergent property, and Q^1 causes the instantiation of Q^2 , another Q-property. However, if all Q properties are emergent from P properties, then Q^2 must itself have a corresponding micro-level P-property, P^2 , from which it emerges. But if Q^2 is emergent from P^2 , then the instantiation of P^2 is sufficient to cause the instantiation of Q^2 , which leaves us with an overdetermination problem: both P^2 and Q^1 are sufficient to cause the instantiation of Q^2 . The overdetermination is systematic; every case of emergence will bring with it an overdetermination problem, which challenges the idea that emergents bring novel causal contributions to the world.¹⁹³

In addition to this general causal exclusion problem, the prospect of *downward* causation generates an additional problem for those who endorse the principle of the completeness of physics.¹⁹⁴ The prospect of downward causation violates the principle of the completeness of physics, because downward causation involves a macro-level property

¹⁹² Kim, J. (2006) pg 557

¹⁹³ This is a rough version of the overdetermination problem generated by downward causation as presented by Kim in Kim, J. (2006).

¹⁹⁴ See pg 25 for details on the principle of the completeness of physics.

causing an effect at the micro-level, and this cannot be reduced to an instance of micro-micro causation.

There are a number of responses to these problems, both the causal exclusion problem and the particular problem of downward causation. One could argue against setting up our understanding of emergence in this way; for instance, I have portrayed emergence as a relation that obtains between many micro-level properties and one macro-level property, whereas this version of the problem of downward causation presents emergence as a one-one relation. One could argue against the idea that property instantiations cause other property instantiations. One could also simply deny the principle of the completeness of physics. There is a lot in this discussion to respond to, but in the rest of this section I will leave much of this discussion to one side. I will argue that, even if we take this presentation of the problem of downward causation on its own terms, the prospect of downward causation does not pose a problem for the perspectival account of emergence.

The response to the problem of downward causation from the perspectival account of emergence is fairly simple: there is very little room in the perspectival account of emergence for causation. In building this account, I did not start with the assumption that emergents must make novel causal contributions. Indeed, I argued in Chapter 4 that this assumption about the concept of emergence is false; the idea that emergent properties have novel causal powers is a feature of some treatments of emergence, but is far from universal. Furthermore, the resulting account presents emergence as relative to an observer, rather than part of the mind-independent universe, and there are few accounts of causation that can reconcile this with the idea that emergent properties are in some sense the locus of the causal efficacy of macro-level entities. Because of its observer-relative status, the perspectival account of

emergence has little to do with causation, and so the problem of downward causation is no problem for the perspectival account of emergence.

Section 6: Objections and Replies

In this section I will consider some objections to the perspectival account of emergence and offer responses to those objections.

6.1 Objection 1: This is simply an epistemic account of emergence

The first objection I will consider is the position that the perspectival account of emergence offers nothing over and above what is offered by a standard epistemic account of emergence. According to this line of thought, in developing the perspectival account I have brought nothing new to debates about emergence, and the appeal to observers adds nothing to the familiar position that emergence merely indicates the limits of our knowledge.

I have a number of responses to this objection. First of all, epistemic accounts of emergence are often presented as dealing with a different sort of phenomenon from those dealt with by metaphysical accounts. Some proponents of epistemic accounts of emergence think that there is an epistemic “type” of emergence, and another, “metaphysical” type too.¹⁹⁵ I, however, am coming from a very different place. I have argued that metaphysical accounts of emergence do not survive the collapse problem. Accordingly, I do not present the perspectival account as a complementary alternative to a metaphysical account of emergence, but as the only game in town.

¹⁹⁵ See, for instance, Chalmers, D. (2006), Wilson, D. (forthcoming)

My second response focuses on differences in the content of a standard epistemic account, as opposed to the perspectival account. The perspectival account, in permitting convergence of observers on cases of emergence and differentiation among cases of emergence based on various factors, permits far more variation and differentiation between cases of emergence than is permitted in a traditional epistemic account of emergence. For instance, Hempel and Oppenheim, who offer an epistemic account of emergence, make no attempt to address questions about *for whom* some property is emergent. The perspectival account, as its name suggests, takes perspective far more seriously. In claiming that some property is emergent, we must specify *who* it is emergent for, and also what standards they adopt for explanation, availability and so on. Traditional epistemic accounts of emergence, on the other hand, don't permit this focus on the particular perspective from which some property is emergent, nor do they permit differences between cases of emergence on the basis of differences between the observers for which the relevant properties are emergent.

Another new element of the perspectival account of emergence, which makes it different from a simple epistemic account of emergence, is the explicit focus on explanation. According to the perspectival account, emergence tracks explanatory relations, rather than the absence of knowledge, a feature that takes it beyond a merely epistemic account of emergence. For these reasons, the perspectival account of emergence is a genuine new alternative rather than an epistemic account in disguise.

6.2 Objection 2: Permits bizarre cases

The perspectival account of emergence permits a wide, varied range of cases of emergence. Consider three cases of emergence: a case that obtains because the relevant

observer has a misguided conception of scientific explanation, a case that obtains because the observer doesn't have access to a particular computer program, and a case that obtains because it is impossible for any human observer to develop an explanation of the relevant sort. The fact that the perspectival account categorizes all of these as emergent might be seen as a problem for the position; perhaps it would be better to acknowledge no emergence at all than to permit such a wide and odd variety of cases.

In response, I would acknowledge that the perspectival account does permit a wide range of cases, and that some of those cases are emergent for seemingly odd, uninteresting reasons. However, I will first of all point out that there is no conceptual requirement that emergence be *rare*. Some metaphysical accounts of emergence present emergence as a rare phenomenon, but this is not a central element of the *concept* of emergence. In many cases, across philosophical and scientific accounts of emergence, emergence is presented as prevalent.¹⁹⁶ The fact that the perspectival account presents emergence as prevalent is not in itself an objection to the account.

The more challenging idea in this area is the idea that emergence is, generally speaking, supposed to be at least *interesting*. If this is true, then the fact that the perspectival account permits cases that are uninteresting is a reason not to endorse it. In response, I would return to the rich resources that are available to differentiate between cases of emergence. The perspectival account does not, as I have emphasized already, involve treating all cases of emergence as equal. Indeed, a central part of the new contribution made by the perspectival account comes in understanding emergence as a common, variegated phenomenon.

¹⁹⁶ See Mitchell, S. (forthcoming), Kelso, J.A. Scott (1995)

Permitting cases of emergence that are emergent for “uninteresting” reasons, such as observers who cannot formulate an explanation because they don’t have the right computer program, isn’t a problem so long as we don’t have to treat all cases of emergence as equal. Endorsing the perspectival account does bring with it a very different way of thinking about emergence overall. It is a shift from thinking of emergence as something homogeneous to something heterogeneous. But the fact that it involves this shift is not an objection to the account, particularly when we consider the resources available to differentiate between cases.

6.3 Objection 3: What about deduction?

I have presented an account of emergence based on explanation. One could argue that this is a conceptually inadequate account of emergence because so many accounts of emergence are based on failures of deducibility.¹⁹⁷ Is this a failure of the perspectival account, that it doesn’t present a failure of deducibility as at least necessary, if not sufficient, for emergence?

In response, I would point out that in Chapter 3 I argued that an account of emergence *must* include the explanation condition in order to avoid the collapse problem. Accounts of emergence based on failures of deducibility face the collapse problem and so aren’t viable. More generally, however, an account based on explanation can encompass cases of emergence that involve deductive failures, depending on the notion of explanation in question. If the observer is looking for an explanation that involves deduction and none is available, then we have a case of emergence that involves a failure of deducibility.

¹⁹⁷ See Mitchell, S. (forthcoming)

The appeal to deducibility, and other, similar features such as unpredictability, reflects the very simple idea that in cases of emergence we constrained our attempts “move from” facts about the micro-level to facts about the macro-level, whether that move is understood in deductive terms or otherwise. The explanation condition for emergence, however, captures this idea perfectly, and so we lose none of the intuitive appeal of the association between emergence and deductive failures when we embrace an account of emergence that involves explanation.

6.4 Objection 4: This isn't about metaphysics

A common response to the perspectival account of emergence runs along the following lines: what this account does is provide us with tools to identify many different kinds of explanatory gaps that obtain for many different reasons. But this account does not say anything about metaphysics and so isn't a complete account of emergence. A complete account of emergence would ideally address questions about the viability of metaphysical non-reductionism.

In response I would point out that the collapse problem and the failure of various responses to it show that there is no viable, non-arbitrary metaphysical conception of emergence. So I have said something about the metaphysics of emergence, in so far as I have argued that emergence is not a metaphysical phenomenon, but instead an explanatory one. Emergence, I have argued, is not the metaphysical basis of non-reductionism.

Of course, the perspectival account of emergence is compatible with both metaphysical non-reductionism and with metaphysical reductionism, so long as neither position is thought to be metaphysically under-written by emergence. But the perspectival

account of emergence presents emergence as a non-metaphysical relation, and so *does* make metaphysical commitments. The perspectival account of emergence is not an incomplete metaphysical account of emergence, but instead a complete non-metaphysical account of emergence.

Section 7: Conclusion

In this chapter I have formulated and defended the perspectival account of emergence, building on the results of Chapters 2 through 4. I have discussed in detail some different features of this account, as well as considering some objections to the view. In the concluding section of this dissertation I will offer some broad reflections on the nature of emergence and on the prospects for a more general perspectival approach to metaphysics.

Summary and Conclusion

I began this discussion of emergence by focusing on two purportedly characteristic features of emergence. Firstly, emergence is a relation that is used to track purported distinctions between metaphysical levels, and secondly, “emergence” is a “*term of art*”, different in the hands of whomever chooses to use it. After working through the collapse problem and developing the perspectival account of emergence, however, we face a very different picture. According to the account of emergence I have developed, emergence is an explanatory, rather than metaphysical, phenomenon and doesn’t track distinctions between metaphysical levels. Furthermore, I have offered a precise definition of emergence, in place of a philosopher’s “*term of art*”.

In these concluding remarks I will summarize the content of Chapters 1 through 5, and will then offer some broad reflections on the nature of emergence according to the perspectival account of emergence. I will also discuss the potential benefits of taking a similar perspectival approach to other debates in metaphysics, beyond the topic of emergence.

In Chapter 1, I introduced the idea of emergence. I showed that emergence is often used to vindicate metaphysical non-reductionism of different kinds. I also explored some other uses of the concept of emergence, including epistemic and computational uses. I then traced the historical development of the concept, through its early incarnations in the work of

the British Emergentists, through a fallow period in the early 20th century during which appeals to emergence weren't prevalent in philosophy, through to the contemporary scene in which the idea of emergence is popularly used by philosophers and scientists. I discussed some related concepts, including reductionism, supervenience and physicalism.

In Chapter 2, I presented a problem for metaphysical accounts of emergence, the *collapse problem*, which indicates that emergence tracks relationships between arbitrary groups of properties rather than distinctions between levels of properties in nature. I explored the idea that the collapse problem involves a kind of cheating, and that in formulating it, I had attributed far too broad a group of properties to the micro-level. I attempted to flesh out this criticism in a number of different ways by restricting the micro-level properties in line with a metaphysical distinction, such as the distinction between intrinsic and extrinsic properties and the distinction between dispositional and non-dispositional properties. In each of these cases, however, these attempts to solve the collapse problem for metaphysical accounts of emergence failed.

In Chapter 3, I offered an alternative solution to the collapse problem. I argued that emergence is commonly associated with the unavailability of certain scientific explanations, and codified this association into a condition for emergence, the explanation condition, according to which a property is emergent if and only a certain scientific explanation is unavailable. The explanation condition, combined with a minimal restriction on scientific explanation, provided a solution to the collapse problem because, given that collapse objections don't provide resources for explanations, they aren't enough to make a given macro-level property non-emergent. Understanding emergence in terms of explanation allowed me to solve the collapse problem.

At this point, however, my account of emergence was not complete. While exploring the connection between emergence and explanation, I observed that emergence is associated with the unavailability of radically different kinds of explanation, including mechanistic explanations, a priori reductive explanations, computational explanations and explanations involving models of certain systems. The collapse problem is a general problem, and so to offer a general solution I realized that I needed to make room for this variation. So I argued that the relevant kind of explanation varies depending on who is conducting the investigation, and so that what it takes for some property to be emergent varies from observer to observer.

In Chapter 4, I offered a rational reconstruction of the concept of emergence based on the solution to the collapse problem that I developed in Chapter 3. Taking Carnap's criteria for success in rational reconstruction – similarity, exactness, fruitfulness, and simplicity – I argued that an explication of the concept of emergence as the unavailability of a certain scientific explanation for an observer met these criteria better than alternative explications. I argued that this explication offered a unified concept of emergence, which, unlike most philosophical treatments of emergence, allowed us to regard philosophical and scientific uses of the concept of emergence as continuous.

Finally, in Chapter 5 I formulated a positive account of emergence, the perspectival account of emergence, according to which a property is emergent iff a certain scientific explanation is unavailable to an observer. I discussed some of the central features of this positive account, and looked into different ways of differentiating between the wide range of cases permitted by the perspectival account of emergence.

The perspectival account of emergence offers a new way of thinking about emergence. According to this view emergence is relative to an observer, among other factors,

and it is common and heterogeneous. There are many different cases of emergence, which obtain for many different reasons. On this account, describing some property as “emergent” is the *beginning* of a conversation, rather than the end. Once we have described a particular property as emergent, we can then move on to ask *why* it is emergent, for *whom*. We can ask whether this property is emergent relative to an alternative notion of explanation, or of unavailability. Furthermore, both philosophers and scientists can conduct research into what makes some property emergent. This makes for not only a deeply variegated notion of emergence, but also one that unifies philosophical and scientific thinking about emergence. This account of emergence enables us to move beyond philosophical debates about cases of emergence that *start* with a conception of emergence as necessarily metaphysical or necessarily epistemic. Instead we have an account of emergence that permits emergence to obtain for many different reasons.

I have argued that emergence is not a metaphysical phenomenon, and that appeals to emergence should not be used in attempts to vindicate metaphysical non-reductionism. However, I have also argued that the perspectival account of emergence offers a useful tool to the metaphysician, because it offers a way to catalogue and taxonomize all of the different types of case of emergence. Adopting this perspectival account of emergence makes precise exactly what work needs to be done in order to argue for the position that some case of emergence obtains for a metaphysical reason. For example, I have argued that the metaphysical non-reductionist would have to argue that a particular observer or observers, a particular notion of explanation, a particular notion of unavailability and a particular conception of the part-whole relation are metaphysically privileged, in order to show that any particular case of emergence obtains for a metaphysical reason.

One may be tempted to interpret this perspectivalism as part of a broader commitment to reductionism, or even relativism about metaphysics, but this is not the case. I have argued that, because of the collapse problem, emergence is not a metaphysical relation. However, I am optimistic about the possibility of metaphysics, conceived of as the delineation of the structure of the mind-independent universe. What's more, I think that a broader perspectival approach to metaphysics, and not just to emergence, could provide a helpful *supplement* to a more traditional approach to metaphysics.

I did not begin this research by *deciding* to understand emergence as a perspectival relation; the need to solve the collapse problem brought me to that result. But in formulating the perspectival account of emergence, it has become clear that a perspectival approach has some interesting benefits that may generalize to other cases. The unique contribution of the perspectival approach to metaphysics is, as the name suggests, that it encourages us take perspective very seriously. Using the perspectival approach to emergence, I was able to provide tools for taxonomizing and specifying in detail what each instance of emergence consists of. This clears the way for the person who wants to argue that one of these cases of emergence is underpinned by some metaphysical relation to argue that one of these explanatory profiles in particular is metaphysically privileged. Similarly, the person who argues that no case of emergence obtains for a metaphysical reason must argue that none of these explanatory profiles can be a guide to metaphysics. In each case the perspectival approach to emergence makes the next step clear, and precisifies exactly what would be required to justify a metaphysical commitment.

A perspectival approach to some other phenomenon could follow this pattern. Before engaging in metaphysical enquiry a philosopher could carefully document explanatory

practices, the objects that they posit for pragmatic reasons, the categories that have a special status because they are pragmatically valuable, and so on. Having done so, the philosopher could survey the results and only *then* identify cases that are ripe for further metaphysical investigation. The idea behind this approach would be to take metaphysical commitment slowly and carefully, and to ensure that when we make a metaphysical commitment, the feature is not in fact a feature of our explanatory practices, our perspectives or our interests, but is instead a genuine feature of the mind-independent universe.

I anticipate that a similar, perspectival approach will be of great use in the study of other phenomena. Accordingly, in future research I intend to develop a general perspectival approach to metaphysics and argue for its usefulness through a number of different case studies, beyond this application to the case of emergence.

Bibliography

- Alexander, S. (1920) *Space, Time and Deity*. Macmillan.
- Anderson, P.W. (1972) “More is Different.” *Science* Vol 177 no. 4047 pp 393-396.
- Armstrong, D. (1978) *Universals and Scientific Realism*. Cambridge University Press.
- Assad & Packard (1992) “Emergence” in Bedau, M. & Humphreys, P. eds (2008) *Emergence*. MIT Press.
- Balazs and Epstein (2009) “Emergent or Just complex?” *Science* Vol. 325 no. 5948 pp. 1632-1634.
- Balog, K. (forthcoming) “In Defense of the Phenomenal Concept Strategy.” *Philosophy and Phenomenological Research* 84 (1):1-23.
- Bechtel, W. & Abrahamsen, A. (2002). *Connectionism and the mind: Parallel processing, dynamics, and evolution in networks*. Blackwell.
- Becker, G. (1962) “Irrational Behavior and Economic Theory”, *Journal of Political Economy* 70: 1-13.
- Beckermann, A. (2000) “The perennial problem of the reductive explainability of Phenomenal consciousness.” in T. Metzinger (ed.) *Neural Correlates of Consciousness – Empirical and Conceptual Questions*. MIT Press.
- Beckermann, A., Flohr, H., and Kim, J., eds. (1992) *Emergence or Reduction?* Walter de Gruyter.
- Bedau, M. (2003) “Downward Causation and Autonomy in Weak Emergence”. In Bedau, M. & Humphreys, P. eds. (2008) *Emergence*. MIT Press.
- Bedau, M. & Humphreys, P. eds. (2008) *Emergence*. MIT Press.
- Benton, E. (1974) “Vitalism in Nineteenth Century Scientific Thought: A Typology and Assessment.” *Studies in History and Philosophy of Science Part A* 5 (1):17-48.
- Bird, A. (2007) “The Regress of Pure Powers?” *Philosophical Quarterly* 57 (229):513–534.
- Blackburn, S. (1990) “Filling in Space.” *Analysis* 50 (2):62-5.

- Broad, C. D. (1925) *The Mind and Its Place in Nature*. Routledge.
- Bromberger, S. (1966) "Why-Questions," in Colodny, R. ed., *Mind and Cosmos*. University of Pittsburgh Press.
- Carnap, R. (1950) *Logical Foundations of Probability*. University of Chicago Press.
- Chalmers, D. (1996) *The Conscious Mind*. Oxford University Press.
- Chalmers, D. (2006) "Strong and Weak Emergence" In Davies & Clayton (eds.) *The Re-Emergence of Emergence*. Oxford University Press.
- Chalmers, D. and Jackson, F. (2001) "Conceptual Analysis and Reductive Explanation" *Philosophical Review*. 110, 315-61.
- Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again*. MIT Press.
- Crane, T. (2001). "The Significance of Emergence." In Carl Gillett & Barry M. Loewer (eds.), *Physicalism and its Discontents*. Cambridge University Press.
- Craver, C. (2007). *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford University Press, Clarendon Press
- Davidson, D. (2001) *Essays on Actions and Events*. Clarendon Press.
- Dennett, D. (1991) "Real Patterns." *Journal of Philosophy* 88 (1):27-51.
- Fodor, J (1974). "Special Sciences." *Synthese* 28 (2):97-115.
- Fruton, J. (2006) *Fermentation: Vital or Chemical Process?* Brill.
- Hanna, J.F. (1968) "An Explication of 'Explication'." *Philosophy of Science* Vol 35 (1):28-44
- Hempel, C. & Oppenheim, P. (1948). "Studies in the Logic of Explanation." *Philosophy of Science* 15 (2):135-175.
- Hempel, C. & Oppenheim, P. (1965) "On the Idea of Emergence." In Bedau & Humphreys ed (2008) *Emergence*. MIT Press.
- Hohwy, J. & Kallestrup, J. (2008) *Being Reduced: New Essays on Reduction, Explanation, and Causation*. Oxford University Press.
- Holland, J. (1999) *Emergence: From Chaos to Order*. Helix.
- Hoover, K. (2001) *Causality in Macroeconomics*. Cambridge University Press.

- Horgan, T (1993) "From Supervenience to Superdupervenience: Meeting the Demands of a Material World." *Mind* 102 (408):555-86.
- Humphreys, P. (2008) "Computational and Conceptual Emergence." *Philosophy of Science* 75 (5):584-594.
- Jackson, F. (1982) "Epiphenomenal Qualia." *Philosophical Quarterly*. 32, 127-136.
- Jackson, F. & Pettit, P. (1992) "In Defense of Explanatory Ecumenism." *Economics and Philosophy* 8(01)
- Johnson, S. (2002) *Emergence: The Connected Lives of Ants, Brains, Cities, and Software*. Scribner.
- Kelso, J.A. Scott (1995) *Dynamic Patterns*. MIT Press.
- Kim, J. (1982) "Psychophysical Supervenience." *Philosophical Studies* 41: 51-70.
- Kim (1995) Entry on "emergent properties" from *The Oxford Companion to Philosophy*, ed Ted Honderich.
- Kim, J. (1999) "Making Sense of Emergence." *Philosophical Studies* 95 (1-2):3-36
- Kim (2003). "Supervenience, Emergence, Realization, Reduction." In Loux, M. & Zimmerman, D. eds., *The Oxford Handbook of Metaphysics*. Oxford University Press.
- Kim, J. (2006) "Emergence: Core Ideas and Issues." *Synthese* 151 (3):547-559.
- Koppel et al (2008) "Experimental Evidence for Spatial Self-Organization and Its Emergent Effects in Mussel Bed Ecosystems." *Science* Vol. 322 no. 5902 pp. 739-742
- Langton, C. (1995) *Artificial Life: An Overview*. MIT Press
- Langton, R. (1998) *Kantian Humility*. Oxford University Press.
- Langton, R. & Lewis, D. (1998) "Defining 'Intrinsic'." *Philosophy and Phenomenological Research*. 58: 333-45.
- Levine, J. (2001) *Purple Haze*. Oxford University Press.
- Lewes, G.H. (1875) *Problems of Life and Mind*, vol. 2. Kegan Paul, Trench, Turbner, and Co.

- Lewis, D. (1983) "New Work for a Theory of Universals." *Australasian Journal of Philosophy* 61:343-377.
- Lewis, D. (1986) *On the Plurality of Worlds*. Blackwell.
- Lewis, D. (2009) "Ramseyan Humility." In Braddon-Mitchell & Nola (eds.) *Conceptual Analysis and Philosophical Naturalism*. MIT Press.
- Manley, Chalmers and Wasserman (2009). *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford University Press.
- MacDonald, C. & MacDonald, G. (eds.) (2010) *Emergence in Mind*. Oxford University Press.
- McKittrick, J. (2003) "A Case for Extrinsic Dispositions." *Australasian Journal of Philosophy* 81 (2):155 – 174.
- McLaughlin, B. (1992) "The Rise and Fall of British Emergentism." In Beckermann, Flohr & Kim (eds.) *Emergence or Reduction?: Prospects for Nonreductive Physicalism*. De Gruyter.
- Merricks, T. (2001) *Objects and Persons*. Clarendon.
- Mitchell, S. (forthcoming) "Emergence: logical, functional and dynamical." *Synthese*
- Molnar, G. (1999) "Are Dispositions reducible?" *Philosophical Quarterly* 50 (194):1-17.
- Mumford, S. (1998) *Dispositions*. Oxford University Press.
- Nagel, E. (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. Harcourt.
- Nagel, T. (1974) "What is it like to be a bat?" *The Philosophical Review* Vol. 83, No. 4
- Ney, A. (2008) "Physicalism as an Attitude." *Philosophical Studies* 138 (1):1 - 15.
- Nitschke, J. (2009) "Systems chemistry: Molecular networks come of age." *Nature* 462, 736-738
- O'Connor, T. and Wong, H. (2005) "The Metaphysics of Emergence." *Noûs*, 39, pp. 658-678.
- O'Connor, T. and Wong, H. (2012) "Emergent Properties", *The Stanford Encyclopedia of Philosophy* (Spring 2012 Edition), Edward N. Zalta (ed.),

forthcoming URL = <<http://plato.stanford.edu/archives/spr2012/entries/properties-emergent/>>

- Oppenheim, P. and Putnam, H. (1958) "The Unity of Science as a Working Hypothesis" in H. Feigl et al. eds. *Minnesota Studies in the Philosophy of Science*, vol. 2. Minnesota University Press.
- Papineau, D. (2011) "What Exactly is the Explanatory Gap?" *Philosophia* 39 (1):5-19.
- Port, R. and Van Gelder, T. (1995) *Mind As Motion: Explorations in the Dynamics of Cognition*. MIT Press.
- Prior, E. Pargeter, R. & Jackson, F. (1982) "Three theses about dispositions." *American Philosophical Quarterly* 19 (3):251-257.
- Purves, D. et al (2008) *Neuroscience (4th Edition)*. Sinauer.
- Quine, W.V.O. (1960) *Word and Object*. MIT Press.
- Redhead, M. (1995) *From physics to metaphysics*. Cambridge University Press.
- Rumbaugh, D. M., Beran, M. J., & Pate, J. L. (2003) "Uncertainty monitoring may promote emergents." *Behavioral and Brain Sciences*. Vol 26, Issue 3, pg 353.
- Salmon, W. (1989) *Four Decades of Scientific Explanation*. Pittsburgh.
- Schaffer, J. (2004) "Two Conceptions of Sparse Properties." *Pacific Philosophical Quarterly* 85 (1):92–102.
- Sellars, W. (1961) "The Language of Theories," in Lange, M. (2007) *Philosophy of Science: An Anthology*. Blackwell.
- Shoemaker (2007). *Physical Realization*. Oxford University Press.
- Sider (forthcoming) *Writing the Book of the World*. Oxford University Press.
- Silberstein and McGeeever (1999). "The Search for Ontological Emergence." *Philosophical Quarterly* 50 (195):182-200.
- Sklar, L. (1967) "Types of Inter-Theoretic Reduction." *British Journal for the Philosophy of Science* 18 (2):109-124.
- Stephan, A. (2006) "The dual role of emergence in philosophy of mind and cognitive science." *Synthese* 151 (3) 485-498.
- Stoljar, D. (2010) *Physicalism*. Routledge.

- Teller, P. (1986) "Relational Holism and Quantum Mechanics." *British Journal for the Philosophy of Science* 37 (1):71-81.
- Van Fraassen, B. (1980). *The Scientific Image*. Oxford University Press.
- Van Fraassen, B. (2004) *The Empirical Stance*. Yale University Press.
- Weatherson, B. (2001) "Intrinsic Properties and Combinatorial Principles." *Philosophy and Phenomenological Research* 63: 365-80.
- Weisberg, M., Needham, P. and Hendry, R. (2011) "Philosophy of Chemistry", The Stanford Encyclopedia of Philosophy (Winter 2011 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/win2011/entries/chemistry/>>.
- Wilson (2010) "Non-reductive Physicalism and Degrees of Freedom." *British Journal for the Philosophy of Science*. 61: 279-311.
- Wilson (forthcoming) "Metaphysical Emergence: Weak and Strong." In Mumford & Tugby eds. (forthcoming) *Metaphysics and Science*.