A Critique of Humean and Anti-Humean Metaphysics of Cause and Law

By Benjamin T. H. Smart, PhD

Thesis Submitted to the University of Nottingham for the Degree of Doctor of Philosophy, May 2012

Abstract

It is my contention that physics and metaphysics (or at least the aspects of metaphysics to be considered in this thesis) broadly strive to achieve common goals: to understand what our physical system is constituted by, and both how, and why it evolves in the way that it does. Metaphysicians, as well as the scientific disciplines, play an important role in our understanding of the universe. In recent years, physicists have focussed on finding accurate mathematical formalisms of the evolution of our physical system - if a metaphysician can uncover the metaphysical underpinnings of these formalisms; that is, why these formalisms seem to consistently map the universe, then our understanding of the world and the things in it is greatly enhanced. Science, then, plays a very important role in our project, as the best scientific formalisms provide us with what we, as metaphysicians, should be trying to interpret – but these interpretations are integral to understanding the nature of natural laws and causation.

In this thesis I examine existing metaphysical views of what a law is (both from a conceptual and from a metaphysical perspective), show how closely causation is linked to laws, and provide a priori arguments for and against each of these positions. Ultimately, I provide an analysis of a number of metaphysics of natural laws and causation, apply these accounts to our best scientific theories, and see how these metaphysics fit in with our concepts of cause and law. Although I do not attempt a definitive metaphysical account myself, I conclude that any successful metaphysic will be a broadly Humean one, and furthermore that given the concepts of cause and law that shall be agreed upon, Humean theories allow for there to be causal sequences and laws (in line with our concepts) in the world.

Acknowledgements

Sections 3.3.7 – 3.4 contain extracts from my 2011 forthcoming paper, with the permission of Philosophical Studies:

Is The Humean Defeated by Induction *Philosophical Studies* (BTH Smart: forthcoming 2011), which has been published online at: http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s11098-011-9767-5. See www.springerlink.com

I would also like to recognise the following institutions for their financial support:

The Bosdet Foundation

The Strasser Foundation

The States of Jersey Department of Education Sport and Culture

My thanks also to Penelope Mackie, Stephen Barker, Michael Talibard, Matthew Tugby, Ben Curtis, Greg Mason, Peter King, Dave Ingram, and the postgraduates of The University of Nottingham, for their invaluable comments; to my mother, and to my stepfather, Ian; and to Charlotte Clarke, without whose love and support this thesis would never have been completed.

Contents

Introduction:

Part 1 The Concept of Cause

Chapter 1 A Conceptual Analysis of Causation

Chapter 2 (I) Three Conceptions of Causation

Chapter 2 (II) The Concept of Natural Laws

Chapter 2 (III) Laws and Causation Explaining Together

Part 2

Chapter 3 Hume's Regularity Theory of Laws of Nature

Chapter 4 Sophisticated Regularity Theory

Chapter 5 Armstrong's Governing Conception of Laws of Nature

Chapter 6 Scientific Essentialism

Chapter 7 Neo-Dispositionalism

Chapter 8 Consistency With Our Best Scientific Theories

Chapter 9 The Case Against Dispositional Essentialism

Chapter 10 Thesis Conclusion

Introduction

Though there is not room here to argue for my every basic tenet, it is worth confessing at this point that this thesis is written by someone who holds that physics and metaphysics (or at least the aspects of metaphysics to be considered in this thesis) broadly strive to achieve common goals: to understand what our physical system is constituted by, and both how, and why it evolves in the way that it does. It seems to me that the primary tools of the scientist are empirical evidence, mathematics, and although this is perhaps less appreciated, imagination - these are fundamental to any great scientific breakthrough. For us, the metaphysicians: imagination, science and *a priori* reasoning form the foundation of our enquiries. I believe that for the metaphysician, reasoning without due consideration of science will almost inevitably lead to unjustified, and probably false conclusions.

Metaphysicians play an important role in our understanding of the universe. In recent years, physicists have focussed on finding accurate mathematical formalisms of the evolution of our physical system - if a metaphysician can uncover the metaphysical underpinnings of these formalisms; that is, *why* these formalisms seem to consistently map the universe, then our understanding of the world and the things in it is greatly enhanced. Science, then, plays a very important role in our project, as the best scientific formalisms provide us with what we, as metaphysicians, should be trying to interpret.

In this thesis I examine existing metaphysical views of what a law *is* (both from a conceptual and from a metaphysical perspective), show how closely causation is linked to laws, and provide *a priori* arguments for and against each of these positions. Ultimately, I aim to provide an analysis of a number of metaphysics of natural laws and causation, apply these accounts to our best scientific theories, and see how these metaphysics fit in with our concepts of cause and law. Although I do not attempt a definitive metaphysical account myself, I conclude that any

successful metaphysic will be a broadly Humean one, and furthermore that given the concepts of cause and law that shall be agreed upon, Humean theories allow for there to be causal sequences and laws (in line with our concepts) in the world. First, though, I need to define and explain a few of the notions central to this thesis.

The aim of the first part of this thesis is to outline a conceptual analysis of causation. Ideally this conceptual side of things should not get entangled too much with the metaphysical analyses – however, if we are to talk about the metaphysics of causation and laws of nature, we need to know what concepts we're trying to make sense of. After all, whether or not there are laws, and whether or not there is causation in the world at all, depends on our concept of cause and law. Imagine, for example, that our conceptual analysis of cause and law leads us to conclude that the properties involved in causal sequences need to be of a certain nature (to have certain second-order properties), but, following our metaphysical analysis, it turns to be metaphysically impossible for any property(ies) to be of this nature – in this scenario there would be no causal sequences. I hold a success theory with respect to causation, by which I mean that ours *is* a world with causation, and much of our causal talk should come out to be true - so in the scenario outlined above, I would contend that something has gone wrong in the conceptual analysis. My conceptual analysis will then provide us with the meanings of the terms 'cause' and 'law', and thus the *de dicto* necessities that need to be satisfied for something in the world to *be* a causal sequence, or a law.

One might compare this contrast between the conceptual and the metaphysical analysis of cause and law to questions concerning the 'notion' of a person (the conceptual questions), and the more metaphysical questions about their 'nature' (are they extended physical objects, etc). Another example being an analysis of the notion of value (concept), contrasted with the metaphysical question of whether there is anything 'in the world' that corresponds to values. We certainly have a notion of causation and a notion of law, the aim of the conceptual analysis is to

consider this notion in detail. The aim of the metaphysical analysis is to take a look at the metaphysics of the evolution of the physical system, and to see how the nature of this system corresponds to our notions (if at all).

One of the most contested questions in the debates scrutinised in this thesis is whether there is 'necessity' in cause and law. Does our *concept* of causation include one event 'necessitating' another? If so, is the belief that there is necessity in the world justified metaphysically? Are the laws themselves necessary (that is, could our laws have been different)? Are there necessary connections between events in *this* world, when there might not have been? These concern different kinds of necessity that are central to both the metaphysical and conceptual analyses, and so before embarking on this project it is important that the different kinds of necessity be spelt out in more detail.

I.1.1 Physical Necessity

I call an event *e* physically necessary in world *w* if and only if its occurrence is logically entailed by the full set of propositions describing *w* immediately prior to *e*, and the full set of true law-statements of *w*. For example, if (i) *w* consists in its entirety of one grain of salt and a pint of water; (ii) in *w* it is a true law-statement that all salt dissolves in water; and (iii) the grain of salt comes into contact with the water, then the grain of salt will dissolve. The event's occurrence is logically entailed and hence physically necessary.

I.1.2 Metaphysical Necessity: (a) Between Events

Necessary connection between distinct events is a metaphysically stronger notion than physical necessity. It requires not only logical entailment, but there to be some, what I will refer to as 'metaphysical glue' holding the world together. Metaphysical glue (if indeed it is there and can be made sense of), is whatever it is *in the world* that is the reason some events occur *in virtue* of others. I call it metaphysical glue because, in a sense, it glues worldly properties together (in the case of electronhood and charge -1, metaphysical glue would be seen as whatever it is that glues these properties together; that is, what makes sure all electrons have charge -1). There are various different proposals as to where the metaphysical glue *is*,

and to what it might be – some believe it exists in the form of extrinsic relations between the properties of particulars, some believe the metaphysical glue is intrinsic to the properties of things. But for my purposes, it is enough to say that one event is metaphysically necessary[1] if there is some metaphysical glue in the world ensuring its occurrence.

I.1.3 Metaphysical Necessity: (b) Between Distinct Existences

Necessary connections between distinct events of course fall into this category, but they only form a subset of this class. There may be a necessary connection between a conjunction of entities (not necessarily an event)/events and a further event. This is relevant because Armstrong's view of laws and causation does not involve necessary connections between distinct events, but does involve necessary connections between distinct existences. In Armstrong's case, the 'existence' I refer to differs from an 'event' in so far as the 'existence' is a conjunction of an event and a 'law-*in*-nature'. But more on this later.

I.1.4 Logical Necessity

A proposition is logically necessary if it is true in every possible world. An event is logically necessary if it occurs in every possible world (the proposition describing it is true in every possible world), and a law is a law with logical necessity if it is a law in every possible world. If proposition *L* is a true law-statement in *w* but not in *w1*, then *L* is not logically necessary.

I.2 The Main Contenders

First: two Apologies

This thesis is of wide scope, and so it is not possible to cover all the material written on each of the metaphysical views I consider. The views I focus on are The Regularity View, The Nomic-Necessitation View (the Armstrongian View), Scientific Essentialismm, and what I will call Neo-

[1] In this thesis, metaphysical necessity is completely detached from logical necessity. An event can be metaphysically necessary without its occurrence in every possible world.

Dispositionalism. Because of space restrictions, it is not possible to examine arguments from all the many philosophers who have contributed to this debate, so I focus on only a few of the main proponents of each view, and their arguments in defence of their position and in opposition to the others. In the metaphysical debate, representing the Regularity View will be David Hume, David Lewis, and Helen Beebee; the Armstrongian view will be largely represented by David Armstrong; Scientific Essentialism by Brian Ellis; and Neo-Dispositionalism by Alexander Bird and Stephen Mumford.

Secondly, in this thesis I assume, from the metaphysical perspective at least, that we live in a deterministic world.

Before beginning a more detailed discussion, I outline the four main views of cause and law to be considered, and the notions central to them.

The Humean Naive Regularity Theory of Causation is devoid of all metaphysical necessity and necessary connections between distinct existences; laws are mere regularities in nature, and cause is reducible to law. **The Humean Sophisticated Regularity Theory (SRT)** is also a view devoid of metaphysical glue, but it has stronger restrictions on which regularities count as laws – the SRT account of causation does not resemble the Regularity View, either, as the SRT offers a counterfactual account of singular-causation (See chapter 3), as opposed to the 'cause as an instance of law' approach that the naive regularity theorists uphold.

This second 'Humean' view is the main opposition for the three 'necessitarian' views to be looked at. I call these three necessitarian views the 'Armstrongian' View, Scientific Essentialism, and Neo-Dispositionalism. All three views are distinctly anti-Humean insofar as they deny that there is nothing more to laws and causation than just patterns of events. However, what they *do* pick out as the metaphysical underpinning of cause and law differ significantly. Crudely, where the Regularity Theorist believes that laws turn out to be regularities of the

form ‘all Xs are Ys’ and that there are no necessary connections in the objects at all, **The Armstrongian View** is that there are necessary connections between distinct existences, and that these are a result of natural necessitation relations between universals, which are in themselves ‘powerless’.

In the Armstrongian view, the all-important necessitation relations hold contingently; that is, the relation may hold between F and G in some worlds but not in others[2]. Where N(F,G) does hold, it governs the universals such that where F is instantiated G will also be instantiated. However, it is possible that we can have two qualitatively identical worlds, but at one world N(F,G) holds and at the other it does not.

The neo-dispositionalist on the other hand, holds that the ‘oomph’ in the world is provided by the intrinsic properties of particulars. For the neo-dispositionalist, all properties in all possible worlds are wholly dispositional, and (For Alexander Bird (Bird 2007), at least) the laws of nature supervene on these properties.

Finally, the Ellisian **Scientific Essentialist** holds a similar thesis to the neo-dispositionalist insofar as he accepts that dispositional properties provide the oomph in the world - but proponents of Ellis’s view also believe there to be an abundance of categorical (non-dispositional) properties. Furthermore, the ‘essence’ of scientific essentialism is that the laws of nature describe the essential properties of natural kinds.

I.3 Properties in the Cause and Law Debates

As might already be apparent, ‘properties’ are central to this entire debate. In fact, the differences between the views can, for the most part, be demonstrated in terms of the nature of the properties taken to exist in this

[2] As most philosophers have focussed on David Armstrong’s work on the DTA view (and due to the lack of space required to provide a fully comprehensive account), my discussion will also focus on Armstrong, for whom a law of nature takes the form N(F,G), where N(F,G) denotes a natural necessitation relation holding between the universals F and G (as a point of interest, N is itself a universal). N(F,G) holds contingently at a world, so where there is a natural necessitation relation between F and G at one world it may not hold at another.

and other possible worlds. I now briefly outline the different types of property we will encounter, but a far more detailed discussion of all these property-types will be provided in the relevant chapters.

I.3.1 Dispositional Properties

Both the neo-dispositionalist and Ellisian Scientific Essentialist have ontologies ripe with dispositional properties. These are taken to be intrinsically 'powerful'. They are instantiated in particulars, and in dispositional ontologies they determine (wholly determine in the case of the neo-dispositionalist) how those particulars behave, so if there is such a thing as necessary connection between distinct existences, the dispositionalists hold that these connections exist (at least partly) in virtue of the dispositional properties of things.

In short, dispositions are properties with at least one determinate manifestation (generally events identifiable by a change in the properties of the particulars involved) and at least one determinate stimulus condition; when the stimulus conditions are met, and in the absence of finks or antidotes (a discussion of which will be provided shortly), that property will (as a matter of metaphysical necessity) manifest. A vase (the classic example, although certainly not a fundamental property) has the dispositional property 'fragility'. Fragility's stimulus conditions are numerous, but one of them is 'hitting with a hammer', and its manifestation is 'smashing'. It follows that whenever a vase is hit with a hammer (in the absence of finks or antidotes) it will, as a matter of metaphysical necessity, smash. The identity of these properties is determined by their manifestations and stimulus conditions; that is, their causal and nomological roles.

One might argue that one can simply *define* 'fragility' as the property that plays the dispositional role it plays, and maintain that properties have their dispositional roles as brute fact - if property F plays the fragility role in the actual world, and the property G plays this role in *w**, them (as we've defined a property's identity as *being* its causal role),

property F in our world and property G in *w** are the same property. However, this would be a matter of *de dicto* necessity. The dispositionalists want dispositional properties to play their dispositional roles as a matter of *de re* necessity, such that there is a fact of the matter about which properties play which dispositional roles independently of our defining them into existence. Brian Ellis (Ellis 2000) considers these properties to be 'natural kinds of property' that determine the 'natural kinds of processes'; Bird and Mumford are not forced into accepting a natural kind ontology, but they consider dispositional properties to be universals (or perhaps perfectly resembling tropes) with their identities fixed by their relations to other dispositional properties. For both Ellis and the neo-dispositionalists then, the same property plays the same dispositional role in every possible world. It follows that in two worlds with the same properties, the same equations of motion will apply. As you can see, dispositional properties have a distinctly modal character.

I.3.2 Categorical Properties

Categorical properties confer no causal powers; that is, they are intrinsically inert. They provide no oomph in the world, and they cannot in themselves determine the evolution of our physical system – taken in isolation they provide no metaphysical necessity. Categorical properties are wholly non-dispositional.

I.3.3 The Natural Necessitation Relation (The N-relation)

The natural necessitation relation is a relation between universals that plays a specific nomological role. Relations and properties, although both universals in the Armstrongian ontology (supposedly the only ontology that includes the N-relation), do differ in some important respects. Properties are typically monadic and non-relational, and relations are polyadic. But properties and relations are also closely related in many respects, not least (for Armstrong, at least) in their status as universals.

The N-relation is what Armstrong takes to be a second-order relation. It holds between two or more universals (say, F and G), and ensures that F and G are linked with metaphysical necessity in any world where N holds between them (This is denoted by N(F,G)). The second order relation, N, is itself a universal. For the time being this is all I shall say on the matter, but making any sense of N outside of the nomoligical role it is supposed to play is no trivial matter. I leave further talk of the N-relation to the more detailed discussions of Armstrongianism later in this thesis.

I.4.1 The Modal Character of the Necessitarian Views

There are two modality features I consider. First there is the notion of necessary connection between distinct existences, which we already know is entirely absent in the Humean theories. This necessity concerns what must happen in some world given the state of affairs given the metaphysical glue – a world at time *t* can be seen as a particular distribution of properties, and if there are necessary connections between distinct existences, at least *some* aspects of property distributions immediately subsequent to *t* are determined. In a wholly deterministic world, the entire property distribution at *t+1* is determined, and so any world with property distribution Pd^1 at *t* will have the same property distribution Pd^2 at *t+1*, but whether or not a world instantiates Pd^1 is contingent.

The second modality feature I consider is the necessity of the laws; that is, whether the laws are themselves metaphysically necessary. If they are, then the laws will be the same across all metaphysically possible worlds, if not, there will be metaphysically possible worlds with different laws. This, however, is work for later chapters. We may now embark on the conceptual analysis of cause and law.

Part I

The Concept of Cause

1.1 A Conceptual Analysis of Causation

The purpose Part I of this thesis is to find the best of the candidates for our concept of causation, discuss these candidates' compatibility and natural pairing with the metaphysical positions outlined in the introduction, and to show how all these concepts are closely linked with laws of nature. The analysis below, amongst other things, provides us with the *de dicto* necessities for sequences to be considered causal, and for statements to be law-statements. The evolution of our physical system may be mind-independent, but the notions of cause and law certainly are not..

There are three main contenders for our concept of causation to be considered. The first is 'non-singular causation', where a cause is deemed to be an instance of a law. This is generally associated with the Humean regularity theory metaphysic. This, I think, is easily dismissed. The second is a form of singular-causation known as the 'conditional' or 'counterfactual' account of causation. Although the counterfactual account is also mentioned by Hume (who seems to mistakenly equate it with non-singular causation), it is strongly associated with David Lewis and J.L. Mackie (see Lewis 1973; Mackie 1974). The final conceptual analysis I will look at will be a dispositional account. This, again, is an account of singular-causation, but is characterised by the actualisation of things' propensities/dispositions rather than by counterfactual conditionals. This links most naturally with the dispositional metaphysics. I will argue that the both the dispositionalist and the 'cause as instance of law' views fail to capture our concept of causation, but that the counterfactual account succeeds

1.2 My Starting Point

I think as good a place to start an analysis of the concept of causation as any is to look at what David Hume saw to be the main conditions for a causal sequence. Hume writes

> The relation of CONTIGUITY (is) essential to that of causation… the second relation… is that of PRIORITY of time in the cause before the effect… There is (also) as NECESSARY CONNEXION to be taken into consideration; and that relation is of much greater importance than any of the other two above-mentioned (*T*, I, I, pp 120-125)

In the following sections, then, I consider both whether causes and effects must stand in certain spatiotemporal relations, the asymmetry of causation that naturally follows from these spatiotemporal considerations, and of course the nature of the necessity relation that seems to be central to our conception of causation. I conclude that neither spatial nor temporal considerations play a necessary role in our concept of causation, and furthermore that contrary to what the necessitarians believe, the necessary connection between events in causal interactions can be satisfied without metaphysical necessity; that is, without any kind of 'metaphysical glue' constraining the evolution of our physical system. If necessity is required in our conceptual account, it can be captured by physical necessity, where effects must follow from their causes if statements describing these events are logically entailed by statements describing the laws of nature, and statements describing the states of affairs at the time of causal interaction[3].

1.3 Spatial Contiguity

One posited necessary condition is that of spatial contiguity, whereby if object involving event *a* causes object involving event *b*, the object involved in *a* has to be touching the object involved in *b* (noting that if we advocate this position, that's not to say we require 'every cause to be contiguous with its effect, but merely that where this is not the case, cause

[3] Note the claim here is a conditional one – *if* necessity is required, it can be captured by physical necessity.

and effect are thought to be joined by a chain of intermediate items…' (Mackie, JL 1974: 19)). This condition is perhaps appealing as it rules out the initially counter-intuitive concept of action at a distance. Nevertheless, I don't think it plays a significant role in our concept of causation, nor need it in a metaphysical account, as despite the fact that we normally associate causation with spatial contiguity, it would seem too restrictive to deny the possibility of causal connections between objects spatially separated. We can imagine Bob getting into a Star Trek-style transporter, pressing a button and arriving twenty miles away from where he started. This may be difficult to acknowledge as possible in our world, but only because it is not something we regularly observe. It is neither necessarily inconsistent with current physical theory[4], nor does action at a distance seem totally at odds with our concept of causation. It certainly seems in this case that Bob's pressing the button in the transporter caused his arrival twenty miles down the road.

1.4 Temporal Priority

Another possible necessary condition is temporal priority - that causes must be temporally prior to their effects. I accept the view that causal relations are generally not static facts, having temporally distinct events as their relata, but the temporal priority condition is questionable. I hold that neither simultaneous nor backwards causation are (at least conceptually) impossible, despite our tendency to envisage causes to be temporally prior to their effects. There are even phenomena in our world that suggest 'simultaneous causation' not only to be merely a conceptual possibility, but regularly occurring.

We know from empirical evidence that no matter how far two quantumly entangled particles are separated, one will have the spin-state +0.5 and the other will have the spin-state -0.5; it follows that when the

[4] Quantum mechanics allows for spontaneous movements of this kind at a quantum level, and if all the quantum particles move at the same time precisely 20 metres down the road simultaneously, to all intents and purposes the entire macroscopic object would have moved 20 metres without having passed through any spatial points in between.

spin state of one particle changes, the other changes simultaneously. One interpretation of quantum mechanics suggests that the observation of an as-yet indeterminate spin-state (a wave function) causes the wave function to collapse, giving it a determinate +0.5 or -0.5 spin-state. As soon as the observed particle has a determinate spin-state the as yet unobserved particle with which it is quantumly entangled will also have a determinate spin-state (the opposite spin-state to that observed). If this interpretation is correct, the observed particle's collapsing causes the as-yet unobserved particle's collapse (the latter counterfactually depends on the former), and so the cause is simultaneous with its effect. The view that, in cases like these, the mere act of observation has this causal influence, is perhaps rather counter-intuitive - and indeed it may well be false. Nonetheless, it does not seem conceptually problematic for changes in an observed particle's state to simultaneously causally affect the unobserved particle's state.

Although we tend to think of effects occurring subsequent to their causes, as we have seen it is conceptually possible for an effect to occur simultaneous to, or even prior to its cause (any story with backwards time travel seems to involve backwards causation). One might respond, however, that we are making an error in choosing 'external' time as the basis from which we judge temporal priority. David Lewis (1976) argues that there are two kinds of time, 'external time' and 'personal time'. The former is time as measured by calendars and clocks in railway stations (the period of external time passed between midnight on the first of January 29 –time A-, and midnight on the first of January 21 –time B-, is precisely one year), and personal time is time as measured by our own watches. Every individual has their own personal time, and although in general, personal time is synchronised with external time, it is possible for less personal time to pass than external time between times A and B - although one year of external time may have passed, it is possible for only two minutes of an individual's personal time to have passed given certain circumstances. Lewis writes that when this happens, when the time of the

journey is unequal to the external time passed, the individual has time travelled[5]. We know that it is physically possible for forwards time travel to occur in this way, but whether or not backwards time travel is possible is still up for debate. However, remember we are not interested in physical possibility for the time being, but in the concept of causation. Fictional novels and films are full of stories involving backwards time travel, so this is clearly a concept we can entertain.

In light of this let us now consider the 'effect must follow the cause' criterion in terms of personal time rather than external time: Bob got in a time machine in 2009 and arrived in 1989. The external time passed is minus twenty years, but the personal time passed is whatever period of time passed on Bob's watch whilst in the time machine. Bob is older in 1989 than he was in 2009, so in terms of Bob's personal time his arrival in 1989 (the effect) is after his departure in 2009 (the cause), and so the 'effect must follow the cause' criterion was satisfied.

Of course now we have to consider what determines the direction and passage of personal time. One common answer to this question is the direction of causation, but if by definition personal time passes in accordance with the direction of causation, trivially all effects are subsequent to their causes in personal time. 'If A caused B then B follows A' becomes trivial. It therefore seems far more plausible that we should judge the direction of causation by external time. If we do this, though, then there are exceptions to the 'temporal priority' rule. We are left with a dilemma – either in a causal sequence, B is trivially always subsequent to A (personal time), or we must dispense with temporal priority as a condition for *a* to cause the cause of *b* (external time). I would suggest the latter is the more intuitive choice.

[5] Time dilation in special relativity shows this to be possible, and empirical evidence has verified the hypothesis. Time travels slower for objects at high velocities, so it is theoretically possible for a man to get in a space rocket, go round and round the Earth very quickly for 10 years, and come back to find 100 earth years have passed. Forward time travel is therefore not only possible, but regularly occurs – even if not by noticeably large amounts.

1.5 The Asymmetry of Causation

It does seem to be the case, however, that there is an asymmetry in causation. When *A* causes *B* it is not also the case, at least not conceptually, that *B* causes *A*. If we were to assume that causes always precede effects in external time this would be easy to deal with. We could simply identify the temporally prior event, *A*, as the cause. However, as I have just shown, there is no conceptual (and perhaps even physical) problem with simultaneous or even backwards causation in this respect. But we still need to account for this asymmetry somehow, so we must look elsewhere.

J.L. Mackie's attempt to capture the asymmetry of causation appeals to 'fixity'. He claims that X cannot be considered causally prior to Y if Y was fixed in the circumstances prior to X's being fixed, where X is fixed if X has already occurred or if circumstances sufficient for X have already occurred. He claims that (a) if X is fixed before Y is fixed then X is causally prior to Y; (b) if Y is fixed as soon as X is fixed, and X is unfixed until it occurs (this will be the case when X is sufficient in the circumstances for Y) then X can be causally prior to Y; and (c) 'if there is some continuous causal process linking X and Y, and if Z was not fixed until it occurred, then X was causally prior to Y'. (Mackie 1974: 190)

Mackie describes causes as events that are sufficient in the circumstances for their effects, and so if X occurs and Y is the effect of X, then there can be no cases where X is fixed and Y is not fixed. Circumstances (a) will therefore never come about. This problem is not so problematic if we discard Mackie's belief that causes must be sufficient in the circumstances for their effects[6]. However, as David Sanford points out there are serious problems with (b) and (c) also. If X is fixed as soon as Y is fixed, then of course Y is also fixed as soon as X is fixed. Y would, under (b), be as causally prior to X as X is to Y. Similar problems arise when we consider (c). If both Y and Z become fixed when X becomes

[6]More on this later

fixed, then X is causally prior to Y and Z. 'According to [b] however, both Y and Z are causally prior to X'. (Sanford 1976: 195)

David Papineau addresses the asymmetry problem in his (1985) paper: *Causal Asymmetry.* He shows that there is a probabilistic link between two effects of a common cause, but not between two causes of a common effect – whence the asymmetry. Obviously central to Papineau's account is the idea of probabilistic asymmetry, but Papineau's account of causal asymmetry still involves the conjunction of causes being, in the circumstances, sufficient for their effects (it's just that the conditions obtaining may not be determinate). Note that given that this section is concerned with the conceptual analysis of causation, it should be acceptable to use probabilities (at least epistemic probabilities) in our analysis, for we do have a concept of probability[7], despite my assumption of a deterministic universe.

We often identify true 'general' causal-statements by their apparent probability-raising attributes. For example, we may well make the following the claim:

"Smoking is a cause of cancer because it raises the probability of getting cancer, but it is not sufficient for getting cancer."

But Papineau shows that one can accept this statement whilst holding on to the claim that whenever cancer is caused, its cause was sufficient in the circumstances to bring it about. Consider the following example[8]:

S = Smoking

X = Some unknown extra condition

A = Inhaling asbestos

C = Cancer

[7] A discussion of the details of our concept of probabilities is not required here. I will leave it as a brute fact that we have a notion of 'X making Y more likely, though Y may still not occur', which is all that's required.

[8] See Papineau 1985 pp273-289

Suppose that S&X are together sufficient for contracting cancer (but S and X are individually insufficient), and that A is sufficient for cancer. Suppose also that one can only get cancer through S&X or A. Then 'if you smoke and had that extra condition' or you inhale asbestos, you will get cancer. AND, if you get cancer, you either smoke and have that extra condition, or you have inhaled asbestos.

$S\&X \vee A \leftrightarrow C$

If S&X obtains then *ex hypothesi* C is determined, and S is an inus condition (an *insufficient* but *non-redundant* part(s) of an *unnecessary* but *sufficient* condition – see section 2.2) of C. But S is also a probability raiser. Prob(A) is necessarily smaller that Prob(XvA) so long as prob(A) < 1 and prob(X) > 0. If one doesn't smoke the Prob(C) = Prob(A), and if one smokes then Prob(C)= Prob(X v A). Smoking raises the probability of getting cancer, and so the general causal statement that smoking causes cancer will be deemed to be true.

So how does this help us with an account of causal asymmetry?

If $S\&X \vee A \leftrightarrow C$ then $C\&\neg A \rightarrow S$

Assume S leaves a trace 'T' – say, yellow fingers. We now have:

$C\&\neg A \vee T \rightarrow S$

As Papineau points out, C is clearly an inus condition S, so there is no asymmetry. How do we deal with this? Consider smoking as a potential cause of both cancer and yellow fingers.

$S\&X \vee A \leftrightarrow C$

$S\&M \vee N \leftrightarrow Y$ (M= not scrubbing fingers) (N=being jaundiced, a chalk worker…)

There is therefore a 'probabilistic association between C and Y'(Papineau 1985: 279). X and A are completely causally isolated from M and N, it is fairly obvious that cancer and yellow fingers are likely to show up

together fairly regularly relative to what might be expected given their independent probabilities of occurring. 'Intuitively smoking makes cancer more likely, and smoking makes yellow fingers more likely, and so cancer makes yellow fingers more likely'(*ibid*). If X has cancer it's likely that X smokes, and smoking gives you yellow fingers.

However, what if we consider the 'inverted inus conditions':

$C\&\neg A \vee T \leftrightarrow S$

$C\&\neg S \vee R \leftrightarrow A$ (where R is the traces of asbestos)

We can't argue that smoking is more likely in asbestos factories, because the inverted inus conditions (¬A, T, ¬S, R) are *not* probabilistically independent of one another. T is negatively correlated with ¬S, and R is negatively correlated with ¬A. These negative correlations cancel out the appearance of C in both C&¬AvT and C&¬SvR and leave us with a null correlation between smoking and asbestos factories (*ibid*).

In other words, in the case of a common cause for two effects, the background conditions for e1 and e2 are independent of one another, so the occurrence of the background conditions for e1 does not affect the background conditions for e2. But we can't say that smoking makes working in an asbestos factory more likely, because although C is a condition of both, the background conditions for S include conditions that negate the background conditions for A. When T obtains above, C&¬S cannot, and where R obtains, C&¬A cannot.

So one can justifiably assert an association between two effects of a common cause, but there is no association between two causes of an effect. The link between the two effects of a common cause enables you to say "e1 is more likely given e2", because the background conditions leading to e1 and e2 from the common cause are probabilistically independent. However, there cannot be shown to be an association between two causes of an effect. In the case provided, the two causes are smoking and working with asbestos, where cancer is the common effect.

So we ask, is there a connection between smoking and working with asbestos? With the inverted inus conditions in the necessary and sufficient conditions for smoking, cancer is conjoined with 'not working in an asbestos factory' in the first disjunct, but of course this is incompatible with 'working with asbestos'. For me, Papineau has clearly demonstrated that an asymmetry between causes and effects is an important part of our concept of causation, and that it can be accommodated within our conceptual analysis.

1.6.1 The 'Necessity' Condition

These first few potential conditions for causation are interesting in their own right, but perhaps the most controversial and interesting condition is the necessity condition – the requirement that for a pair of events to be considered cause and effect, there must be a necessary connection between the cause and the effect.

I claim in chapter 2 that if the thought that one event necessitates another is a part of our concept of causation, the necessity is of the physical necessity variety – that is, necessary given the laws of nature, as opposed to necessary given the metaphysical glue. However, it is nonetheless important to consider the *nature* of metaphysical necessity, as it has been claimed that metaphysical necessity between causes and effects is an unintelligible idea.

David Hume (1739-4) explicitly states (and I believe quite in line with our intuitions) that the most important aspect of our concept of causation is the necessary connection between causes and effects. We have to consider both what 'necessary connections' are, and if and when they can be applied to object-involving events.

Hume writes in the Enquiry that:

> It is universally allowed, that matter, in all its operations, is actuated by a necessary force, and that every natural effect is so precisely determined by the energy of its cause, that no other effect, in such particular circumstances, could possibly have resulted from it (*E*, VIII, I, p149)

and in the Treatise:

> (With causation) there is a NECESSARY CONNEXION to be taken into consideration; and that relation is of much greater importance (than the relations of contiguity and precedence in time) (*T*, I, II, p125)

Hume sees necessary connection as an integral part of our concept of causation, but traditionally he is interpreted as denying the existence of necessity 'in the objects'. This can be seen just by looking at his 'rules by which to judge of cause and effect' (*T*, I, XV) in the *Treatise*, which do not include a necessity condition of any kind - it seems that for Hume, then, necessity plays a role in how we think about causation, but not in the metaphysics of causation. This brings forth the worry that, given my assumption of an success theory with respect to causation, Hume's conceptual analysis might be flawed (if there is no causation in the metaphysics of causation, there would be no causation in the world. But there is causation in the world!). But as we will see, given Hume's interpretation of what the idea of 'necessary connection' corresponds to in the world, this worry turns out not to be problematic after all. So what does Hume mean by necessary connection? Two questions arise:

1. According to Hume, where is the necessity if it is not in the objects? and;

2. If necessity is not in the objects, would this other 'kind' of necessity satisfy the necessary connection condition in our concept of causation? In other words, must the necessary connection condition that needs to be satisfied for an event to be considered causal be 'in the objects', or can it be located somewhere else?

J.L. Mackie (Mackie 1974: 12-13) makes a distinction between two types of necessity: the first he defines as 'whatever is the distinguishing feature of causal as opposed to non-causal sequences', and the second provides 'the supposed warrant for an *a priori* inference'. As we shall see, Hume shows through his copy principle that an idea of the latter cannot be

obtained - we would need to 'have an impression[9] in between our impression of [the cause] and out impression of [the effect] distinct from both these impressions, that is an impression barely of the producing' (Beebee: 2006: 78). Hume makes the move from this conclusion, to the denial that we have an idea of first kind of necessity, and this, it seems, might be unwarranted. The assumption Hume makes, as Beebee states, is that 'the detection of causation would have to be a detection of something that would generate inference from causes to effects' (Beebee: 2006: 79). One might argue that Hume's assumption here is unjustified. We might well detect a connection between the cause and effect which renders the events to be not entirely 'loose and separate', despite our inability to directly observe the metaphysical glue. However, it seems clear to me that we *do* detect some connection between cause and effect (albeit perhaps not the metaphysical glue). At this point I move on to Hume's arguments as to (a) why we cannot detect metaphysical glue, and (b) to what he believes the idea of necessary connection to correspond to.

To answer these questions from Hume's perspective, a more detailed account of his conception of causation and necessary connection is required, and to provide this we first need to examine his method of reasoning.

Hume held that there are two fundamental kinds of perception, impressions and ideas. The former are vivid perceptions that directly appear to the mind: we have an impression of an apple when we directly perceive the apple, and an impression of a sound when listening to it. Ideas, on the other hand, are less vivid, and are copied from the impressions. An idea of an apple could be a memory of a previous experience, or perhaps a product of the imagination, or its appearance in a dream. Importantly, whenever we have an idea, it has always been copied from at least one impression (the copy principle). Of course we can imagine red unicorns, or trees that sing, neither of which we will ever have directly perceived (so our imaginings are not direct copies), but

[9] A vivid, direct perception.

nevertheless, these complex ideas will have been copied from a conjunction of impressions of red things, horses, horns, trees, and sounds. Even if the shape is entirely new, we could not have conceived of it without impressions of other shapes. If we accept the copy principle, then, somebody who has always been blind cannot form the idea of a red apple, and somebody who has never had a sense of smell cannot imagine what an apple pie would smell like. Ideas that have no corresponding impressions are unreliable, and although they may seem to somehow correspond to reality, in truth the terms we use to refer to them are contentless.

Hume claims that '(we never perceive) any quality which binds the effect to the cause and renders the one an infallible consequence of the other. We only find that one does, in fact, follow the other'. There is no direct impression of necessary connection. But given the argument above the question becomes: "if we can't have a *direct* impression of necessary connection from which to derive the idea, is there any other sort of impression that could enable us to form the idea of necessary connection, and what would this impression be like?"

If we accept Hume's empiricist intuitions, then when we observe a causal interaction all we see is one event followed by another. When a brick hits a window, we have an impression of the brick hitting the window and an impression of the window smashing. There is no further object of which we have an impression from which the idea of necessary connection can be derived. In fact, it is difficult to imagine what this impression would be like.

If there is no impression from which the idea can be derived, then the idea of necessary connection is 'contentless'. To say that there is a necessary connection between Xs and Ys is, in Hume's words, 'unintelligible', which (arguably) is akin to saying "there is a bleurgh between Xs and Ys". But Hume does not make this claim (and we should hope not!). He writes writes that 'necessary connexion betwixt causes and effects is the foundation of our inference from one to the other' (*T*, 1,

XIV), so he certainly believes we have an idea of necessary connection. In fact, he expressly ranked it as the most important aspect of our concept of causation - so where does this idea come from?

1.6.2 The Mind 'Spreads Itself' on to the World

Hume suggests that the impression of necessary connection is an 'impression of reflection' resulting from the observed constant conjunction of events of one type, A, and events of another type, B. Our minds have this propensity, when having observed a constant conjunction between As and Bs, to expect an event of type B when observing an event of type A. When this occurs an impression of the mind is produced, and it is this impression from which the idea of necessary connection is derived.

> (The impression of) necessity, then, is the effect of this observation, and is nothing but an internal impression of the mind, or a determination to carry our thoughts from one object to another. (*T*, I, XIV, p215)

But what is the nature of this feeling? It cannot be the mental event of forming the idea of an event of type B from the impression of an event of type A, as just like in the physical realm, this would simply be a case of two distinct events. Harold Noonan writes that 'the sentiment or impression Hume refers to can… only be an accompaniment to the transition from the idea of an A to the idea of a B, perhaps a feeling of helplessness or inevitability that occurs in the mind when the disposition to make the transition from an idea of an A to the idea of a B is actuated' (Noonan 2007: 83). It is from this feeling that the idea of necessary connection is derived.

We then, it seems, instinctively and mistakenly attribute necessary connections to the objects – according to Noonan, Basson (see Basson 1958: 66-67; Noonan 2007: 84) suggests that for Hume, to attribute necessity to objects is tantamount to attributing emotions to a specific sound (sounds obviously don't have emotions!). If we consider a note despondent, the despondency is in actual fact an emotion in the listener, triggered by the note. The thought is that we begin to attribute emotions to

specific sounds through association: the noise made by a siren in war-time is considered despondent because it incites feelings of fear and sadness, but the emotion is in the mind, not in the note itself (see Basson 1958: 66-7).

Is it possible that through our mind's propensity to (in Hume's words) 'spread itself onto the world', we unwittingly attribute the feeling of inevitability (the impression from which the idea of necessary connection is copied) to the objects themselves? According to this interpretation of Hume, then, we copy the idea of necessary connection from an impression of reflection, which must turn out to be something like a feeling of inevitability (which is itself caused by the observation of constant conjunctions) - we then make a grave category error: we mistakenly believe that the idea of necessary connection has been copied from an impression of sensation (of something in the world), and so inadvertently attribute our feeling of inevitability on to the objects; that is, we actually unknowingly attribute emotions to inanimate objects.

1.6.3 Beebee's Projectivist Interpretation

Alternatively, one could claim that when we say X causes Y (and in so doing assign a necessary connection between X and Y), we are really just expressing our inference from the observation of event X to the occurrence of event Y – we are certainly not attributing emotions to the objects in the sense described above. Beebee believes we should give such a projectivist account of Hume's analysis of causation, whereby causal talk 'is non-representational or non-descriptive: in speaking and thinking causally, we *express*[10] our habits of inference and project them on to the world'. I find this interpretation more appealing than the frightening thought that we attribute emotions to inanimate objects (every time we identify a causal relation!), not least because the previous interpretation renders all out causal talk false, and this does not seem to be Hume's intention.

[10] My emphasis

Regardless of whether one infers Beebee's projectivist, or Noonan-style conclusions from Hume's arguments, ultimately the impression of necessary connection is not an impression of metaphysical glue located in the external world. 'The efficacy or energy of causes is neither plac'd in the causes themselves, nor in the deity, nor in the concurrence of these two principles; but belongs entirely to the soul… Tis here that the real power of causes is plac'd, along with their connection and necessity' (*T*, I, XIV, p216). Whatever necessary connection is, for Hume it is in the mind, not in the objects.

1.7 Galen Strawson: We Can Refer to Necessary Connections

If the Humean concept of necessity is right, then it doesn't *look* to be the same kind of necessity many intuitively believe there to be in causal sequences - the Humean necessary connections do not seem to be plausible candidates for the necessary connections we need, as we really require some objective metaphysical glue between events.

Is it possible though, that by 'Y has to happen if X happens', all we really *can* mean is that 'there is a necessary connection between X and Y', in the sense of necessary connection that Hume suggests? The term 'necessary connection' would refer, but it would refer to something in the mind (perhaps we are referring to the emotion we experience after observing Xs).

Prima facie this might seem implausible, but if it were the case, the Humean concept of necessary connection would be the only kind of necessary connection we *can* refer to, and hence the kind of necessary connection required from the start. But this, one can't help but feel, is a little unsatisfying. The question that needs answering, then, is whether the Humean is right in thinking we cannot refer to anything 'in the objects', from which it follows that certain events have to follow certain other events.

Although what Hume himself believes is unimportant for the purposes of this discussion, Galen Strawson has a different interpretation of Hume that provides us with an alternative conceptual analysis of causation; one in which there really is necessity in the objects. He argues that Hume's claims about necessity being in the mind are purely his 'epistemological' beliefs; that is, he supposes that Hume was simply showing that we can never know the true nature of power or necessary connections. That is not to say, Strawson argues, that there really is nothing in the world like power or necessary connection. In fact he believes quite the opposite is true, it's just that we know nothing about their nature. Nonetheless, the regular succession of objects does indeed totally depend on these powers and forces whose nature we are so utterly ignorant. (Strawson 1989: 279).

I think Strawson is wrong to interpret Hume in this way, but what is interesting is that Strawson believes it is possible to uphold the copy principle and refer to parts of reality of which we have no 'positively contentful or descriptively contentful conception' (Strawson 1989: 278), including that of necessary connection. It is simply not the case that from Hume's theory of meaning we are required to reject the existence of anything of which we cannot know the true nature, or 'real essence'. He argues that when Hume describes something as 'unintelligible' he does not use the term in the same way as we tend to; that is, for Hume, even if a term is unintelligible, that term may still refer to something. According to Hume's theory of ideas we are not able to form an idea of necessary connection with any positive content; we may never be able to know the true nature, or essence of necessary connection, but (according to Strawson at least) that does not mean our term 'necessary connection' cannot refer to something in the world. In the same way as when I claim that "something, let's call it X, smashed my window", it does not seem implausible to say that I can refer to 'X' without knowing (or even being able to know) its nature, and furthermore claim that "necessary connections are the reason ('something', an aspect of reality of which we

can form no idea with positive content) why certain events must follow certain other events".

This looks inconsistent with Hume's position. When we refer to necessary connection we cannot be referring to something that's both 'in the objects' and 'in the mind', and as we have seen Hume clearly states that necessary connections are in the mind - this must follow from his conclusions regarding the impression from which the idea of necessary connection is copied, but if these conclusions are correct, how can the term 'necessary connection' refer to anything in the objects? If, when I attribute a necessary connection between two electrons repelling one another, I'm really just expressing the 'idea' corresponding to the accompanying feeling of inevitability that would be experienced when observing two electrons in close proximity to one another (or expressing a feeling of inevitability), I cannot be referring to some metaphysical glue out there in the world.

Despite the considerable differences between Strawson's interpretation and what I take to be Hume's views on causation, Strawson's theory (taken as a theory of causation in isolation as opposed to a reading of Hume) is compelling. If we can refer to necessary connections, even if we cannot know their nature, we can still claim that there is a real reason why certain events have to follow certain other events, and the counter-intuitive notion that any event can follow any other turns out to be false.

I have issues with Strawson's argument, though. His example seems plausible. We may well look at the dent in a car and presume that '*something* caused it, I know not what'. But is it really true that we have no positively contentful idea of what that thing is like? We have ideas about things that make dents in cars by having impressions of observables making dents in other observables. Although we may not know precisely what caused the dent, we can infer that it was 'caused'; that is, we can conclude that some event occurred, the type of which is constantly

conjoined with the denting of cars. Another car hitting my car could have caused the dent; a brick being thrown at it could have caused the dent, and so forth. Every suitable event we can think of will involve some kind of moving object with mass, and although we don't know precisely what this object would have looked like, we do know something about its nature. We have positively contentful ideas of the kind of thing that cause dents, and we know that were we in the right place at the right time, we would be able to identify *exactly* what massive object it was. We can refer to this 'something', because we have some kind of idea of what this thing may be. The idea of the 'something', in this case, has (at least some) positive content.

But this is not parallel to the case of necessary connection. It is *not* possible, according to Hume's copy principle, to observe metaphysical glue in the objects, even if we are in the right place at the right time. The idea of necessary connection (in the metaphysical glue, sense) thus has no positive content whatsoever, and so the analogy, on the face of it at least, seems to fail.

Chapter 2

In chapter 1 I considered some of the main contenders for inclusion in our concept of causation, and an asymmetry between cause an effect and necessary connection between them seemed paramount. The necessity talked about in chapter 1, however, differs from the necessity I shall consider in this chapter. Here I discuss conceptions in which the necessity is that associated with lawhood. It can be seen not as a conceptual phenomenon, but more in terms of entailment – that is, what circumstances are sufficient for an effect, given certain premises. In this chapter, I first look at a number of conceptual analyses in more detail: the possibility of causes being instances of laws, and two accounts of singular causation. I then look at our concept of what it is to be a law of nature, discussing the possibility of *ceteris paribus* laws, and whether or not laws must govern there instances. I finish with a more detailed account of necessity in causation, and the roles played by laws and causation in explanation.

Part I: Three Conceptions of Causation

2.1 Causation as an Instance of Law

Hume, when considering the nature of the cause-effect relation, saw a cause to be an instance of a constant conjunction. If we assume that all laws are exceptionless regularities (clearly a non-trivial assumption!), then perhaps all causes are indeed instances of laws, but cause being an instance of law is at best a necessary condition. It is by no means sufficient. Let us assume that it is a law that all electrons have charge -1. It is nonetheless fairly obvious that having charge -1 is not an instance of causation. However, as we shall see, laws do play a significant role in

causation. The next question to ask is whether one must know the laws if one is to identify a cause-effect relation that (in some sense) depends on the truth of that law.

It seems to me that it is possible that there are worlds in which the inhabitants do not know what the laws are (and indeed this is probably true of our world!), but they do have a concept of causation (in our sense of the term). These inhabitants can identify a rock hitting a slab of ice as the cause of the ice smashing without knowing what the laws are. The inhabitants may have a concept of law, and they may even believe that each causal event is an instance of a law, but surely they need not know what the true law-statements are in order to identify an instance of causation. However, for Hume, to know something to be a causal sequence is already to know that it's a law.

The concept of cause cannot simply be that a cause is an instance of law, and that the laws do not need to be known to identify a causal interaction. This, I think, leads to the conclusion that one should be looking more specifically at singular causation; that is, rather than looking at more general types of causal interactions and picking out instances of those types as causal interactions, we should be looking purely for an account of causation capable of picking out individual causal events without reference to event-types. The conditional analysis looked at in the next section requires no knowledge of what the laws actually are, only the notion of logical entailment from law-statements.

2.2 J L Mackie and the Conditional Analysis

It seems plausible to say that "if A caused B, then both A and B occurred, and (in the circumstances) if A had not occurred then B would not have occurred… (where A and B are logically and conceptually distinct events)". The brick caused the window to smash if the both the brick hit the window and the window smashed, and if the brick hadn't hit the window the window would not have smashed-. This can be explained in terms of possible worlds by considering a non-A world with previously

exact match of particular facts to the actual world, and let this possible world 'run on' in accordance with the laws of our world - '…the (counterfactual conditional) analysis requires that (the worlds) evolve thereafter in accordance with the actual laws' (Lewis 1983: 43). If B does not occur, then this second counterfactual condition is satisfied. As Lewis writes, '…we think of a cause as something that makes a difference, and the difference it makes must be a difference from what would have happened without it' (Lewis 1973: 557).

I look at Lewis's account in more detail later, but this general way of looking at singular causation is what J.L. Mackie terms 'the conditional analysis'. Mackie provides a comprehensive discussion of the position in *The Cement of the Universe*, a summary of which I provide below[11].

When an Ace (effect 'Z') is served in a tennis match, one may determine the cause to be player A's hitting the ball very hard with good technique. However this couldn't be seen as the only factor involved. The opponent, player B, also plays a role (in not reaching the ball in time). Player B's lack of pace clearly contributes to the resultant Ace, so effect Z should really be seen as a conjunction of causes: XY - Cause 'X' being player A's hitting the ball hard, and cause 'Y' being player B's slow running. X is the event we would identify as the cause, but it is not sufficient for Z. To emphasise, XY (where XY represents the conjunction of causes X and Y) may well be sufficient for Z, but it is not the case that X is sufficient for Z, alone.

Neither is it the case that all Aces require particularly hard serves of the kind denoted by X. A softer serve with a lot of spin (Q), for example, combined with the slow pace of player B may also result in Z. X is therefore neither a necessary, nor sufficient condition for Z, but nonetheless we identify it as the cause.

To avoid the conclusion that sufficiency is not a relevant factor in picking out a cause, one option is to introduce a weaker form of

[11] Using my own examples

sufficiency, whereby we fix the circumstances such that X 'in the circumstances' is sufficient for Z. In the case above the circumstances would include player B's lack of pace, and so X would, in the circumstances, be sufficient for Z.

Can we now claim that if X is sufficient for Z in the circumstances then it must be the cause of Z? - sadly not. It is easy to construct situations where a condition is sufficient in the circumstances but is clearly not the cause of the effect. Suppose that event A is Bob stubbing his toe on a rock, and event B is Bob's dying. Now suppose the circumstances are such that just as Bob stubs his toe a tidal wave crashes onto the beach where poor Bob is located. A was not causally relevant to B despite A's being sufficient for B in the circumstances. This weaker sense of sufficiency looks to be far too inclusive[12].

Mackie thus suggests we introduce the notion of 'sufficiency in the strong sense', whereby we keep the 'in the circumstances' criterion but introduce a new counterfactual conditional: '(in the circumstances) if Y had not been going to occur, X would not have occurred' (Mackie, J.L. 1974: 39). In the tidal wave sequence above, sufficiency in the strong sense would not apply to Bob's stubbing his toe, and so we get the welcome result of it not being a cause of his death. However, *in the circumstances*, if player A had not served an ace, then player A would not have hit the ball so hard, and so his good serve still counts as the cause.

Mackie refers to events like X as *inus* conditions: '*insufficient* but *non-redundant* part(s) of an *unnecessary* but *sufficient* condition' (Mackie 1974: 62) There could, of course, be numerous *inus* conditions that can be associated with the same effect, so when we consider what the cause of an Ace is (on a more general scale) we are left with a disjunction of conjunctions – Z results from (XY or QY or 'etc'). Assuming that 'etc' represents a finite number of conjunctive conditions, we may assert that Z

[12] It will turn out that it is examples of this kind that show why laws alone are insufficient for picking out causal events, but I will address this issue in greater detail later on.

is always preceded by (XY or QY or ‘etc’), and conversely that all (XY or QY or ‘etc’) are followed by Z.

To take this further, the effect Z is also partially reliant on the *absence* of certain occurrences; a gust of wind (D), for example, may force the tennis ball out of play. Perhaps these considerations should also be incorporated into the concept. We are then left with the rather more complicated ‘Z iff (XYnot-D or QYnot-D or...). Clearly the exhaustive list of conditions that need to be satisfied for most causal sequences of this kind would be enormous. Nevertheless, Mill writes that ‘the cause…philosophically speaking, is the sum total of the conditions positive and negative’ (Mackie, J.L. 1974: 63, quoting Mill (1911: Book 3, Ch.5, Sect 3)) But, as Mackie points out, ‘if we go so far as to say

This ‘full cause’ is obviously not what we, as observers, would normally associate with the concept of cause. We identify specific events, or at most relatively small combinations of events as causes, as opposed to these extensive disjunctions of conjunctions. When we identify causes, we are almost always referring to the *inus* conditions within this full cause.

Looking back to whether or not causes necessitate their effects it seems that although X may not itself necessitate Z, the conjunction of X with certain other conditions does. But the counterfactual ‘if X had not occurred then Z would not have occurred’ would often be false, as there are numerous other conjunctions of events that might lead to the same effect. If we consider only the conjunctive condition that *actually* causes Z, however, if we omit X from that conjunctive condition it would not have been sufficient. In some sense at least it looks as though X is a necessary condition, in that the conjunctive condition without X would not be sufficient for Z. To put it another way: if none of the other minimally sufficient conditions for Z are satisfied, then X is both individually necessary, and in the circumstances (whereby all the other conjuncts of the relevant conjunctive condition are satisfied) sufficient for Z.

To see how this fits nicely into the conditional analysis of causation, consider the following example: a foul is committed in a football match. The referee 'sees the offense' (X) and 'blows the whistle for a free kick' (Y). We conclude that the referee blowing his whistle was caused by his observation of the foul, and claim the counterfactual 'if the referee hadn't seen the foul he would not have blown his whistle' to be true. However the referee has assistants on the touchlines. In this case, the assistant saw the foul and raised his flag (R) to indicate this to the referee. Thus regardless of whether or not the referee saw the foul himself, he would still have blown his whistle. The counterfactual is therefore false. However, if we now consider the 'full cause' to be P (which is a disjunction of conjunctions comprising every possible combination of factors that could give rise to Y) as proposed by Mill, evidently 'if not-P then not-Y' is always true, thus satisfying the counterfactual conditional. Identifying P as a cause using counterfactuals does not seem to be troublesome, but this is trivially the case. We're interested in successfully identifying the *inus* condition(s), X, as a cause, rather than P.

We know that if X occurs, but none of the other minimally sufficient conditions (ie R) within the full cause are satisfied then Y also occurs. If we now consider a similar instance, except on this occasion X does not occur either, then it would follow that '(Y) did not occur in F[13]; such an inference justifies the assertion within the scope of the supposition that he (did not see the foul), that Y did not occur, that is, that (he did not blow his whistle), hence sustains the conditional (if he had not seen the foul, he would not have blown his whistle)'[14] (Mackie, J.L. 1974:65).

2.3.1 Singular Causation and Dispositions

In later chapters I discuss the 'metaphysical' aspects of the dispositional analysis of causation in detail (whether or not dispositional properties are natural properties of objects that provide the causal oomph, or metaphysical glue that determines the evolution of physical systems). In

[13] 'F' refers to the background conditions under which a causal event takes place
[14] Bracketed areas indicate my substitution of Mackie's example for my own.

this section I merely wish to introduce the concept of objects being disposed to act in certain ways, how this would fit into an account of singular causation, and ultimately how I believe even a dispositional analysis is closely connected with laws of nature.

Markus Schrenk (forthcoming) summarises Lewis's definition of dispositions as follows:

> Something x is disposed at time t to give response r to stimulus s, iff x has some intrinsic property B so that: if x were to undergo stimulus s at time t and retain property B long enough, s and x's having of B would cause x's giving response r. (Schrenk cf. Lewis 1997: 157)

Under this account, an event would be identified as causal if it involves the manifestation of at least one of the object's dispositional properties B when the stimulus conditions for that disposition are met.

The dispositional analysis thus supposes that objects have certain properties, the manifestations of which constitute the causal events and the evolution of the physical system. This position in the modern debate is taken to be distinctly anti-Humean, yet conceptually speaking, objects can be seen to be disposed to act in certain ways regardless of whether one accepts an anti-Humean view of causation. 'Fragility' is often given as a classic example of a dispositional property, but on the face of it an object being fragile could be interpreted in terms of counterfactuals, without any commitments regarding whether or not fragility is a natural property providing causal oomph. One could conceptually define an object's possessing dispositional properties in the following way:

Disposition B is initially defined in terms of its stimulus conditions and its manifestation – 'fragility' is the dispositional property such that s is 'being hit with a hammer' and r is 'smashing'. An object x possesses dispositional property B iff, if the stimulus conditions for property B are met then property B is manifested' – An object is fragile iff, if that object were to be hit with a hammer it would smash. So (again purely on the face of it) a conceptual analysis of an object being disposed to B only requires

one to consider how that object would behave under certain circumstances.

This counterfactual analysis of dispositions runs into trouble, however, when considering finks and antidotes. For example, a poisonous snake's venom looks to be disposed to kill those who have the venom injected into their blood-stream via the snake biting its victim – under the account presented above, this disposition could be captured by the counterfactual 'if person *x* had been bitten, *x* would have died'. However, if the requisite antidote is taken in time the disposition of the venom to kill will fail to manifest. The counterfactual 'if *x* had been bitten, *x* would have died' would be false, and so the venom would be determined not to have that dispositional property. Yet this is clearly false. Similarly, a negatively charged object, *a*, is disposed to move towards a positively charged object, *b*, if caught in its electro-magnetic field. But if that negatively charged object is also within the gravitational field of a separate larger massive object, *c*, *a* may not move at all (if the two forces exerted are equal), or may even move away from *b* if the gravitational force exceeds the electro-magnetic force exerted on *a*.

Perhaps these issues can be cleared up by giving a more detailed account of the stimulus conditions that includes absences, rather like in the complete disjunction of conjunctions proposed by Mackie. I discuss these issues further in the chapter on the metaphysical analysis of dispositions.

2.3.2 But is Necessity Necessary?

In these final remarks concerning the conditional/counterfactual analyses, I should like to point out that the necessity required (in Mackie's case, the necessity to which I refer is seen in terms of 'sufficient in the circumstances') is thought by some to be absent in many cases of causal interaction. Markus Schrenk (forthcoming in Noûs) argues that even when considering deterministic causation 'whenever a process, starting with event C and finishing with event E, is temporally extended, that is, whenever E is supposed to succeed C after a time Δt, there is the in

principle possibility of an interference with C such that E could have been prevented'. This looks bad not only for those who believe that some form of metaphysical glue provides *de re* necessity between events, but for any conception of causation requiring physical necessity as a necessary condition of causal interactions (which under my interpretation includes all the previously outlined metaphysical views of causation).

I don't believe Schrenk's objection is too problematic when we consider what will be my approach to the kind of necessity required for a causal interaction. It may well be the case that interfering factors during Δt may prevent the effect event we usually associate with a cause event from occurring, but this just amounts to 'if the circumstances had been different' A might not have caused B. I do not deny this. As we saw with Mackie, the circumstances, or background conditions, are vitally important – and these background conditions are contained within the propositions describing the state of affairs at the time of the cause. I use Mackie's *inus* condition analysis to pick out 'the cause'.

Schrenk suggests that that dispositions provide more than contingency because they 'tend towards' their manifestation, but less than necessity 'for only in a derived sense do dispositions, when triggered by complete world states, 'necessitate' their manifestations', but even if he is correct, it seems to me that only this derived sense is required to satisfy the necessity condition in our concept of causation, whatever metaphysical view we adopt. As we shall see in chapter 7, according to Bird's (2007) neo-dispositionalism the laws of nature supervene on the dispositional properties of objects in the physical system, so for him, any causal interactions that take place do so in virtue of the laws. Given the laws and the distribution of the property instantiations at *t*, any causal effect observed at *t*+1 happens as a matter of physical necessity - for any causal interaction, it is still the case that the state of affairs at *t*, plus the laws of nature, necessitate the effects at *t*+1.

Chapter 2, Part II: The Concept of Natural Laws

2.4.1 The Concept of Natural Laws

As with the concept of causation, I consider the task of explicating the concept of law to be separate from the task of investigating what laws in nature *are* (what is their nature), and indeed, in Stephen Mumford's words, whether 'laws can exist *in nature itself*[15]' (Mumford 2004: 9) at all. These questions are unquestionably metaphysical in that one's conclusions will affect one's worldly ontology. The conceptual analysis, on the other hand, has no ontological implications. There may be no laws *in* nature, but we certainly have a *concept* of law; after all, we could all sit down and list many statements we consider either to be, or at least to express. However, as with the concept of cause, the outcome of the conceptual question may well have important implications when we begin to look at questions regarding laws in later chapters (*de dicto* necessities need to be taken seriously, too).

To begin with I take a brief look at the semantic structure of law-statements, concluding that we should not restrict ourselves to one particular structure, or else we risk ruling out numerous generally accepted law-statements. I then consider a number of widely accepted conditions of being a law-statement – if a statement is either to *be*, or to *express* a law, it must, I believe, satisfy these conditions. After outlining these less contentious conditions, however, I present a number of possibilities that will be endorsed by some philosophers, but not others. In particular I look at the proposals that laws must govern their instances, and that there can be *ceteris paribus* laws. Frequently these disagreements in conceptual analysis play a large role in the corresponding metaphysical debates; as we shall see, there are numerous examples of philosophers arguing that "this metaphysical view of laws cannot work because laws would then not have

[15] My emphasis

property ‘X’, and X is essential to laws” – but for those who deny that X is essential to laws, such an argument is unsound.

2.4.2 The Semantics of Law-Statements

It is not a true law-statement in our world that the force of gravity between two objects is inversely proportional to the cube of the distance between them. This is trivially so because the statement is false, but perhaps it *might* have been a true law-statement, as it certainly seems to be of an acceptable semantic form for this status.

Law-statements can take a variety of semantic forms. Laws of nature (one subset of which is ‘causal’) generally seem to simply link certain properties, so one possibility for the semantics is the universally quantified conditional $\forall x(Fx \rightarrow Gx)$, roughly translatable as ‘all Fs are Gs’. Many widely accepted laws of nature (‘all electrons have charge -1’, for example) take this semantic form[16].

Not all law-statements are expressible in this way, however. In this thesis I am also concerned with laws closely connected with causal interactions. With causal laws, the relata can be object-involving events or processes - many law-statements express diachronic laws, so we must also allow statements of the form ‘all Fs are followed by Gs’, *where Fs and Gs are object involving events* to be law-statements too.

But these two semantic-forms still don’t cover all the law-statements we allow, as often the more fundamental (higher-order) laws tend to be functional laws involving forces (where the less fundamental diachronic laws can be derived from, and arguably explained by the more fundamental). Nevertheless these fundamental laws can generally be expressed in the similar semantic form of ‘Fs stand in a functional relationship to Gs’[17]. Consider accelerating bodies where the net force exerted on an object is equal to its mass multiplied by its rate of acceleration. If F is the rate of acceleration of an object of a particular

[16] Whether this is a true law of nature is, nevertheless, debatable.
[17] The kind of functional relationship must be specified of course.

mass, and G is the force acting on that object, then the statement 'Fs (different rates of acceleration of objects of that particular mass) are directly proportional to Gs (different magnitudes of force acting on the objects)' looks to be a true law-statement.

Law-statements, then, can take a variety of semantic forms: some assert the properties certain objects must have; some express diachronic laws, and some express the functional relationship between object-involving events and forces.[18] It seems that specifying precise semantic conditions would rule out many propositions that we accept as expressing laws, and so perhaps it is best not to stipulate exact requirements for semantic structure.

Candidates for Necessary Conditions of Law-Statements

2.5.1 One thing common to all the semantic forms suggested above is 'universality'. Intuitively propositions like 'all objects *in this room* have a gravitational pull between them inversely proportional to the square of the distance between them' are not law-like. Law-statements tend to be non-local; that is, they tend not to refer to individuals, or specify groups of individuals. 'All ravens are black' may be a law of nature, whereas 'all ravens called Ben are black' could not be (see Mumford 2007: 43). Similarly, law-statements are never temporally restricted. However, there is extensive literature concerning *ceteris paribus* laws; that is, laws that allow of exceptions. I will discuss *ceteris paribus* laws in more detail below.

2.5.2 Most philosophers suggest that law-statements, if they are to express laws, must express something with explanatory value and predictive power (or if the laws are the propositions, perhaps they should *be* something with explanatory and predictive power). Laws must either explain, or at least play an important role in explaining some phenomena in the world, and as a result allow us to make predictions about future states of affairs – the

[18] The existence of forces, as a rather more abstract notion than that of massive objects, is a matter of contention. Even if we deny the actual existence of forces, however, they provide a useful mathematical tool that aids expressing uniformities.

law of conservation of momentum explains the behaviour of 'Newton's balls', and allows us to predict what will happen when massive objects collide in the future. I take this to be absolutely of the essence of laws of nature (note though, that having explanatory and predictive power is not sufficient to be a law; Causes, for example, can explain, and understanding what causes what can also help us to make predictions).

2.5.3 Laws are also usually seen to be objective. In Van Fraassen's words: 'Whether or not something is a law is entirely independent of our knowledge, belief, state of opinion, interests, or any other sort of epistemological or pragmatic factor' (van Fraassen 1989: 36). Most would think the Law of Gravitation is a law irrespective of whether or not anyone is around to observe or make inferences about the behaviour of massive objects.

2.5.4 Law-statements should avoid expressing tautologies and other analytic truths; the laws expressed should be discovered empirically, so propositions like 'all bachelors are unmarried men' cannot be considered law-statements.

2.5.5 The true law-statements must be consistent with one another – this is just an assumption I am willing to make.

2.6 Properties in Laws

Laws involve properties – of the metaphysical positions I discuss, laws either describe, supervene upon, or simply *are* relations between properties. It is important, therefore, to have an idea of what a property *is*. The metaphysical work to be done is extensive, and this will be the main focus of all the metaphysical analyses. Conceptually, though, I think we can see properties as qualities we can attribute to things in our world that make for similarities and differences with other things in our world. If two things share all the same properties, then they will be qualitatively identical. If they share very few, they will be very different indeed. I assume for the rest of this thesis that the attributes I list below are central to our concept of properties and property-terms:

1. Each property term refers to a unique quality
2. Two objects that share a property X are qualitatively similar in the X-respect (Two objects instantiating the property 'redness', both look red.)
3. Two objects that share all the same properties aside from spatiotemporal properties are qualitatively identical, but distinct.
4. Properties can play causal/nomological roles in the world
5. Properties are objective features of the world, independent of our identification of them.

The vocabulary used in law-statements should refer purely to natural properties[19]. What makes a property a natural property is contentious – but whatever conception of natural property one adopts, roughly the same properties end up being the natural ones. Often, naturalness is seen to come in degrees: the property 'fragility' might be a natural property, but it is not as natural as the property 'mass'. This raises the question: "how natural must a property be for it to be included in law-statements?" The answer to this question depends, I think, on whether one is willing to accept a hierarchy of laws, and perhaps also on whether one is willing to allow for *ceteris paribus* laws. Again, these issues shall be addressed in greater detail in subsequent chapters.

2.7 Must Laws Govern (Extrinsically)?

One of the more contentious debates about the role laws are supposed to play is that concerning whether or not laws must *govern* their instances. Firstly though, I must explain what I mean by 'govern' in this context.

There are a few metaphysical interpretations of what a governing law might be, but Armstrong sees a world full of intrinsically inert objects, with extrinsic governing laws that affect these objects in such a way that they move around in accordance a set of rules. Imagine governing laws as 'magical invisible fairies with a rule book' - they push and pull the objects in the system around, working in perfect harmony with one another by

[19] For a discussion of natural properties see chapter 5

following the rule-books (for some the rules may differ in different worlds).

However, just as with causation, on the face of it, all we see are regularities. We certainly don't see any magical fairies with rule-books. A discussion of the nature of governing laws will be provided in later chapters, but conceptually, a governing law is extrinsic to objects, and ensures that these objects move in the way that they do. The governing conception is therefore one option for those who want that elusive metaphysical glue[20].

This contrasts with the non-governing conceptions of laws, in which there are neither magical invisible fairies, nor anything playing the role the fairies are supposed to play. Laws, in the case of non-governing laws, are generally seen as a special set of propositions that best describe the world – what's in it, and how these things behave. In this sense, the laws depend on the evolution of the physical system, rather than the other way around.

The issues with the non-governing conception of laws arise elsewhere. Many claim that if laws are just propositions that describe, then they cannot play the other roles laws are supposed to. For example: statements, it is argued, could never play any significant role in explanation, but playing a role in explaining phenomena is a fundamental requirement of being a law. I shall not spend any time here discussing these objections, but in the next chapter I provide a number of arguments in response. Statements corresponding to states of affairs can, I argue, play an explanatory role.

The issue of whether laws govern is of great importance in a metaphysical analysis of laws. But our concept of law does not obviously include a governing aspect, neither does it obviously exclude a governing

[20] Note that what does the governing need not be anything in addition to the natural properties – Alexander Bird, for example, argues that laws supervene on first-order natural properties.

aspect. In light of this, any *de dicto* arguments against there being laws in virtue of there being no 'governing entity', shall not be admitted.

2.8 *Ceteris Paribus* Laws

The concept of law, I claimed, includes their universality; that is, genuine laws hold at all times and all places. But when we actually look at specific examples of laws, this basic assumption looks imprecise at best. In the special sciences, there are plenty of laws that require *ceteris paribus* clauses; that is, they require a little addendum to the effect of: 'if all else is held constant'. In economics, for example, there are 'laws' concerning supply and demand. Keep the supply of a commodity constant, increase demand, and keep all else constant (no corruption etc), then the price of that commodity will rise. This thesis, however, is about laws of nature, and the law of supply and demand is not a law of nature. Our concept of laws in the special/social sciences might allow for *ceteris paribus* laws, whilst our concept of fundamental laws of nature does not.

Many philosophers, however, argue that there are such laws in nature - in fact, some go so far as to say all or nearly all our laws require *ceteris paribus* clauses (Cartwright 1995: 155). But does our *concept* of law allow for exceptions?

There are two distinct ways of interpreting 'exceptions' I wish to spell out, here:

The first is that there can be laws that do not hold universally: 'energy is always conserved' could be a law even if there are occasional instances of energy not being conserved. It is not obvious, some claim, that one exception should prevent a statement from being a law - consider a Humean world where Einstein's equations correctly described the world for the whole of eternity, except at one spacetime point. Should we really rule Einstein's equations out as laws because of this one counter-example? Despite the initial temptation to 'let that one exception fly', it seems to me

that as universality is such a well-entrenched condition of lawhood, we should not take Einstein's equations to be laws in this world.

The second allows for laws to hold universally by introducing a *ceteris paribus* clause into the law itself – there are thus no exceptions to the law, only exceptions to what the law would be without the *ceteris paribus* clause. Those philosophers that argue all laws are *ceteris paribus* in this respect tend to use examples like the forces between an electron and a proton, and how they will move towards one another - because there are nearly always interfering factors (other charged objects in the physical system etc), our predictions about where they will be at certain times can be incorrect if we use purely the laws of electromagnetism, despite these being universally accepted law of nature. However, it seems to me that this is not a genuine *ceteris paribus* law. The law of electromagnetism always holds, even without the *ceteris paribus* clause - our *predictions* are only incorrect insofar as we have failed to factor in other universally holding laws and states of affairs. The laws do not correspond to the *actual* movements/accelerations/charges of things, but to what might be described as the 'accelaratory vectorial contributions' those laws make. The law of electromagnetism gives us a constant, fixed vectorial contribution for the movement of an object. But counteracting vectorial contributions from other laws also contribute to the motion of an object. It might be a law that two protons repel one another, but protons also have mass, and are thus subject to gravitational forces from other massive objects. Suppose there is an uncharged massive object in between two protons placed in such a way that the attractive gravitational force cancels out the repulsive electromagnetic force. In this case, the vectorial contributions to the protons' motion made by the charge is precisely the same as if the massive object was not present, but the vectorial contribution of the massive object of the same strength but opposite direction means the protons remain motionless. When it comes to the fundamental laws, all 'exceptions' seem to be of this kind – but these are not really exceptions at all

We might claim that our concept of laws involves a hierarchy of fundamentality: the laws of physics are more fundamental than the laws of biology. The laws of biology or economics may admit of exceptions, but it seems to me that at the most fundamental level, the laws always contribute the same in any situation. In this thesis I regularly use non-fundamental examples to make my arguments clear, but ultimately I assume that our concept of *fundamental* laws of nature rules out *ceteris paribus* laws; fundamental laws must not allow for exceptions, as this is part of our concept of a fundamental law. I therefore take it to be a *de dicto* necessary truth that there are no fundamental *ceteris paribus* laws.

Chapter 2, Part III: Laws and Causation Explaining Together

In this final part of my conceptual analysis of cause and law I discuss certain areas where I believe that our concepts of cause and law come together. In particular, I firstly claim that all the conceptual (and indeed, it will turn out, metaphysical) analyses of causation are in some way dependent on laws. Secondly, I demonstrate why I believe that both cause and law play a role in our explanations of events and property instantiations.

2.9.1 Physical or Metaphysical Necessity in our Concept of Causation?

There are those who believe that *real* laws must be extrinsic governing aspects of reality: that laws provide the metaphysical necessity in the world, and that, in a sense, laws determine what causes what; there are those who believe in necessary connections 'in the objects' - that the metaphysical necessity derives from the intrinsic properties of particulars[21]; and there are also those who believe that there is no metaphysical necessity whatsoever, but there are laws and causal events

[21] Bird's believes there can be governing laws that supervene on the intrinsic properties of things, but more on this in chapter 7.

nonetheless. The question I want to ask now is: given that there are necessary connections in our concept of causation, what kind of necessity, if any, is required?

In the conditional account of causation provided by Mackie, he talks of causes being necessary for their effects. But Mackie does not ask any questions concerning whether effects follow necessarily from their causes, nor, in particular, what the nature of this 'necessity' would be. For X to be a cause of Y, must Y occur as a matter of logical, metaphysical, or logical necessity (or indeed, with any kind of necessity at all)? It is my intention in this section, to show that our concept of causation, if it requires necessity at all, only requires physical necessity.

There is a clear dichotomy between those who think there is metaphysical glue 'constraining' nature, and those who think there is just one discrete event followed by another, but the arguments concerning whether or not there is metaphysical glue in the world are distinct from the arguments concerning whether or not there is metaphysical glue in the concept of causation. There is, as we shall see, some kind of necessity involved in our concepts of cause and law, but what kind of necessity is involved is not a trivial matter. There are three main options:

1. Physical necessary connections
2. Metaphysical necessary connections
3. No necessity is required

I think nearly everyone agrees that our concept of causation involves some kind of necessary connection between causes and effects (or at least between causes and some 'tendency' to produce an effect), so it is reasonable to rule out option (3); if our concept requires events to satisfy (2), then, *de dicto,* only the anti-Humean ontologies could ever have causation – in which case if we do live in a Humean world, it would be a world without causation (this will be more upsetting to some than others!);

if only (1) is required, then there can be causation in Humean as well as anti-Humean worlds.

The necessitarians will no-doubt jump to conclusion number (2) – that there is something out there in the world that 'glues' certain properties or events together. But for David Hume shows us, it's far from obvious that we can even give the term 'necessary connection in the objects' ('metaphysical necessary connections') any positive content. If the concept of causation (i) includes necessary connections, (ii) we accept that we have no intelligible idea of necessary connections in the objects, and (iii) we disavow Hume's view that we have the kind of idea of necessary connection required (that is, we do not accept that necessity in the mind is sufficient), then the concept of causation is incoherent – it has a necessary condition of which we have no contentful idea. But our concept of causation is not incoherent. Hume provides us with an account of necessary connections that deals with this problem: the necessary connection between distinct existences is really just the *belief* that the effect will inevitably occur given the cause, or a projection of our habits of inference, induced by the observation of constant conjunctions between similar events.

One believes that 'circumstances Y is inevitable given circumstances X (or one projects our habits of inference onto the objects) if one believes that, given circumstances X, circumstances Y had to occur', or in other words: 'all Xs are followed by Ys' - leaving out the more specific modal features of this claim for the time being, it should become apparent that 'inevitability', or the projection of a habit of inference (as I interpret it), and thus the idea of necessary connection, is closely connected to the concept of law, in so far as both are tied to universally quantified statements. If one believes that all Xs are followed by Ys is a law, then one believes that Y is inevitable given X. As we shall see in the next chapter, for the naïve regularity theorist, this translates into: one believes that 'Y inevitably follows X' if one believes that 'all Xs are followed by Ys' is a law of nature:

(a) Y inevitably follows X if X causes Y

(b) Causes are instances of laws

(c) All Xs are followed by Ys is a law

Hume concluded that the idea of necessary connections *in the objects* is unintelligible, and as I show in chapter 3, he concludes that causation is just a matter of regularities in nature. The necessitarians think there is something in the world that makes event Y *have* to follow event X in a metaphysically meaty sense – a metaphysical glue. I believe, though, that Hume was right in thinking the necessity required in our *concept* of causation is linked with thoughts of inevitability, and furthermore that that inevitability is linked to the notion of law.

Physical or Metaphysic Necessity?

Argument 1:

Firstly I must consider what comes first: our concept of law, or our concept of causation; in other words, which of these concepts is 'conceptually prior'?

Given that I accept something like Mackie's *inus* conditions conceptual account of causation, we must have a concept of counterfactuals in order to make sense of our concept of causation. Although I accept that a child may understand that if she hadn't fallen down she wouldn't have felt pain, I nonetheless think the counterfactual analysis of causation involves an implicit reliance on the notion of law. In accepting the conditional analysis, it seems to me that in the context of our conception of causation, whether one event is inevitable, or necessitated by another, can be seen as determined by the states of affairs at the time of the proposed cause of the event, and the true law-statements of the actual world. Furthermore I believe that both the Humean and the necessitarian can accept this conclusion. All we need to get our necessity, then, are

basic law-statements and some rules of inference – no metaphysical glue is required, only physical necessity.

Argument 2:

If our concept of causation requires there to be metaphysical necessity in the world, then a Humean world is a world without causation. Of course, most inhabitants of a Humean world qualitatively similar to ours (that is, one with apparent universal regularities) would have a concept of causation, and they would all believe there to be causation in their world, but they would all be wrong.

For me this is an unacceptable conclusion. When conducting a conceptual analysis of causation, we start by identifying causal events, and then discovering what features link them all together. It is possible that after a general analysis, certain events that were considered causal before are no longer considered causal – we might discover that the effect fails to satisfy the *inus* conditions requirements, for example. In the case of causation, I will assume my success theory is justified. In some cases, a conceptual analysis of X may well lead us to conclude that there are no Xs (we have a concept of fairies, and yet we can conclude that there are no fairies). However, if we were to come up with a conceptual account of causation that, following a metaphysical analysis, lead us to conclude that there was no causation in the world, then I believe something must have gone wrong with our analysis. It might be argued that metaphysical necessity is required for causation, but if one concluded that we live in a world without metaphysical glue, then one would be forced to concede that there is no causation. It is, I believe, *de dicto* necessary that ours is a world with causation, not least because causal roles are *de dicto* necessary conditions for so many of the things we identify as kinds, or as individuals.

Nonetheless, it is also *de dicto* necessary that a world with causation is a world with necessary connections (in one sense or another). To overcome this apparent difficulty, we can claim that only physical

necessity is required. We do not need to know what laws are metaphysically, all we really need to get our necessity are some basic law-statements and some rules of inference. The laws of nature are expressible as propositions – the truthmakers of these propositions will differ depending on the metaphysical view of laws one endorses, but all parties should, I venture, come up with the same list of propositions as expressing (or being, in the case of those who feel laws just *are* those propositions) laws. I propose that the necessity in causation is a matter of logical entailment from propositions expressing states of affairs at the time of causal interaction, and the propositions expressing the laws of nature - the necessity required in causation is thus physical necessity alone, and this need not make any assumptions with respect to whether or not there is any metaphysical glue in the world.

In the next section I show in more detail how physical laws are suited to playing the necessity required in our concept of causation.

2.9.2 The Conceptual Role Played by Physical Necessity

Suppose that there is a causal law at *w* captured by the law-statement "all As are (immediately) followed by Bs", and that at time *t*, A obtains. Naturally it follows through simple logical entailment that B must (immediately) obtain. B necessarily follows given A and the laws of nature, and is thus physically necessitated. If we assume that both the laws of nature, and the causally relevant state of affairs that obtain at any particular time can be captured entirely through carefully structured sentences, then any time one event is necessitated by another there will be a logical entailment similar to that above. If the entailment is not there, then we were wrong in supposing a causal connection between the events. An effect is therefore physically necessary in virtue of the laws of nature holding at *w* and the state of the physical system *w* at *t*, if and only if there is logical entailment from statements representing the relevant aspects of *w*. Of course the language these statements are written in need not be English, indeed generally mathematics is the language of the sciences.

However, it remains the case if a subsequent state X is physically necessitated, then X is entailed by the laws (as mathematical functions) and the initial state of the system (expressed mathematically).

This view seems to correspond with David Lewis's conception of physical necessity. Van Fraassen outlines Lewis's position as follows:

World y is physically possible relative to world x exactly if the laws of x are all true in world y.

It is physically necessary that A is true in world x if and only if A is true in every possible world which is physically possible relative to x.

[Therefore] it is physically necessary that A is true in world x if and only if A is implied by the laws of x. (van Fraassen 1989: 44)

Van Fraassen finds this position somewhat unsatisfactory, however. He writes that '…it is hard to escape the feeling that if the criterion [for physical necessity] can be satisfactorily met in this way, then it must be devoid of all probative force. Doesn't Lewis meet the criterion by robbing it of significance?' (van Fraassen 1989: 45). He thinks that the laws need to explain their instances and why something or some event is physically necessary, 'and no fact can explain anything to which it is definitionally equivalent'. (*ibid*) It seems to me that the intuition that this semantic account does not capture physical necessity probably arises because physical necessity holds between events, not propositions.

But consider for the time being that for each statement expressing a state of affairs (denote this statement: O(*r*)), there is a corresponding state of affairs *r*, where O(*r*) is true at all and every world where *r* is the state of affairs. O(*r*) is thus the proposition that *r* is the state of affairs. We can then discover what events are physically necessitated by taking O(*r*), and deriving the logical entailments from these propositions and the law-

statements[22]. Semantic entailment from statements corresponding to laws and to the state of a physical system at a particular time is therefore at the very least co-extensive with the effects as necessitated by 'real physical necessity', whatever that may be.

These propositions have truthmakers, and although the Humeans might take the truthmakers to take the form of regularities, those holding alternative views might take the truthmakers to be metaphysically real natural necessitation relations, or the genuinely powerful dispositional properties of objects. Given that the holders of the various metaphysical conceptions of causation will have different truthmakers for their law-statements, it seems to me that whatever metaphysical conception you hold, this account of necessity in causation could potentially be appealing to anyone.

What I have said so far only ascertains when an effect is necessitated. I have said nothing about picking out the individual causes of an event whatsoever. The list of propositions O*(r1),* O*(r2),* O*(r3),*...O*(r^n)* – which includes both the law-statements and propositions corresponding to the state of affairs at t - is analogous to Mackie's 'in the circumstances' as discussed in chapter 1. If an event O*(x)* is sufficient in the circumstances for a causally subsequent event O*(y),* then under my account as set out thus far then *x* seems to contribute to the necessitation of *y*, and of course this need not be the case. Stating that the laws and the state of affairs must together entail the effect tells us when an effect is necessitated, but it does not tell us which of the propositions correspond to the actual causes of the effects. To discover these we have to take a closer look at which of the

[22] This section is inspired and partly paraphrased from the relationship between counterfactual dependence among events and their corresponding propositions in Lewis (Lewis: 1973). Lewis writes 'To any possible event e, there corresponds a proposition O(e) that holds in all and only those worlds where e occurs. This O(e) is the proposition that e occurs… Counterfactual dependence among events is simply the corresponding counterfactual dependence among the corresponding propositions.' (Lewis 1973: 186)

propositions representing states of affairs at *t* are causally relevant, and given my account, this will mean selecting which events contribute to the logical entailment. Consider the example provided at the beginning again but with the extra proposition 'X occurs at t'. So now we have

Law 1:	All As are followed by Bs
State of affairs 1:	A occurs
State of affairs 2:	X occurs
Conclusion:	B occurs

Although X is included in the state of affairs at *t* in *w*, it does not contribute to the entailment of B, and thus cannot be considered causally relevant. Remember how in section 2.2 the unfortunate Bob died when struck by a tidal wave. The propositions corresponding to the total state of affairs at *t* and the law-statements would have together semantically entailed Bob's death, but Bob stubbing his toe at *t* also featured in the states of affairs, and so, in the circumstances, Bob stubbing his toe may look to contribute to his death (as it was, in the circumstances, sufficient), when of course it doesn't. We could remove the propositions corresponding to the toe-stubbing altogether, and as with 'State of affairs 2' above, there would be no difference to the result. This then, is the mark of causal relevance in my account. Once we have the total list of propositions O*(r1),* O*(r2),* O*(r3),*...O*(r^n)* we can identify the entailment, and then remove any propositions that do not alter the conclusion. The aspects of reality corresponding to these propositions were not causally relevant.

The same may apply to laws, of course. There may be causally redundant laws included in the set of laws applying at any world when considering one particular causal event. However, as science has progressed the number of fundamental laws has decreased, and it may well end up being the case that one single fundamental law could be applied to any state of affairs and provide the right entailments in any

singular case. Whether this will turn out to be the case or not is a moot point, however, so long as the existing set of laws (even if unnecessarily large) correctly describes the evolution of the physical system. Causes can only be states of affairs (whether they be diachronic or synchronic), the laws are constants that help us pick out which of these states of affairs are causally relevant.

There may be complications, however. Consider the bullet that killed John F Kennedy (JFK). Now suppose I determine (fairly reasonably!) the presence of that bullet in JFK's brain at *t* to be causally relevant to his death. If the proposition corresponding to the presence of this bullet is O*(r1)*, clearly I can't remove O*(r1)* from my set of propositions and retain the semantic entailment of JFK's death at *t*. However, now suppose I go more fine grained than this, and provide every minute local particular matter of fact with its own proposition. There would be a proposition corresponding to the presence of every microscopic particle in that bullet. O*(r1)* would just be shorthand for all the propositions corresponding to all the microscopic particles in the bullet. Let us suppose O*(r1)* is the proposition representing the conjunction of propositions: O*(r1*)*, O*(r1**)*, O*(r1***)*,... (the O*(r1)* list), where this list contains a billion members. Of course, individually, any member of this list could be removed and JFK's death would still be entailed. Indeed, many thousands could be removed at a time and JFK's death would still be entailed. But they couldn't all (or mostly) be removed at the same time, as this would be to remove the bullet altogether, which we have already said would result in the state of affairs expressed as propositions not entailing JFK's death. So which of the propositions the O*(r1)* list correspond to the causally relevant aspects of reality? If we split O*(r1)* into two propositions, O(r^x) and O(r^y), where the former relates to the left hand side of the bullet, and the latter to the right hand side. Removing one of the two would not affect the final result of JFK's death, as either one is sufficient for this effect. We have a classic case of overdetermination, and so following Mackie, I just have to conclude that

both are causally relevant. Despite the aspects of reality corresponding to the propositions O*(r1*)* etc being insufficient for Bob's death, many smaller lists taken from the O*(r1)* list refer to aspects of reality that are. There is a clear resemblance to Mackie's cases of overdetermination here, and so I must conclude that every member of the O*(r1)* list corresponds to a causally relevant state of affairs.

Perhaps it should be mentioned that of course there will be many propositions representing states of affairs that really are relevant in the entailment of the effect-event, many of which we may not generally consider to be the causes. Although this will be the case, we should refer to the discussion provided by Mackie. Any of these propositions will correspond to one conjunct of the conjunction of causes sufficient for an effect. Any individual conjunct can correctly be identified as an *inus* condition.

In summary, if metaphysical necessity is required then only anti-Humean worlds would be worlds with causation. Clearly we (as human beings) have a concept of causation, and if (a) our concept of causation requires metaphysical necessity, and (b) our world *is* a Humean world, then ours is a world without causation. However if physical necessity is all that's needed, a broadly Humean conception (worlds without metaphysical glue) could still capture causation perfectly well. It seems to me firstly that even if we knew this was a Humean world, we would still have a concept of causation that has application to it, and secondly that we can achieve this by requiring only physical necessity.

2.10 Laws and Causation in Explanation

When we are asked to explain something (and let us stick to events for the time being), or when we are asked *why* a particular event occurred, we generally know what kind of things to look out for in order to provide a response, and what kind of responses are expected. If I found my car's front windscreen smashed, I could conjure up a number of possible explanations: "somebody hit it with a baseball bat"; "somebody threw a

brick at it"; "a cricket ball hit the window", etc. These are, at least at first sight, what I would call 'direct causal explanations' – where a direct explanation is one whereby given the explanans and the laws, the explanandum had to occur[23]. Each of these direct explanations could (in principle) be perfectly good explanations of the state of my car.

But these *direct* causal explanations are not the only explanations I might have given. Consider the following: "I stupidly left a 50 pound note on the dashboard"; "I stupidly left the car outside a nightclub in Brixton"; "I stupidly parked next to a cricket pitch". These are also possible explanations, but my parking next to a cricket pitch plus the laws did not necessitate my windscreen smashing, so it is not a direct causal explanation. It is, though, a causal explanation nonetheless. It was *because* I left a 50 pound note on the dashboard that the thief took a baseball bat to it... Both "the windscreen was hit by a baseball bat" and "I stupidly left a 50 pound note on the dashboard" are acceptable explanations, so we must recognise already that (a) not all explanations are *direct* causal explanations, and (b) there can be more than one explanation for an event.

It might be tempting to say that all events causally relevant to event X can serve as explanations for X, but from an intuitive perspective at least, this does not seem right. If I was asked "why was your windscreen smashed", and I responded, "Because two million years ago my ancestor, Bobby, ate 500 grams of oysters", I would no doubt get a strange look. But if Bobby eating those oysters had not caused his feeling an aphrodisiac effect, which in turn caused his sexual activities that night, and so on, then I would not have been around 2 million years later to leave the 50 pound note on the front windscreen.

Here, though, I think we need to make a distinction between two different ways of interpreting 'explanation'. The first way of interpreting

[23] Of course, one can hit a window softly and it would not smash, or one could had bullet-proof glass and it would not smash, or there could be other finks/antidotes – suffice to say, a direct explanation is one where the explanan plus the laws necessitate the explanandum. As a matter of fact, there may be very genuine 'direct causal explanations'.

'explanation' is to see it as an act – a response to a question, either written, spoken, using sign-language, or any other means of communication to whomever asked the *why*-question. The second way of interpreting 'explanation' is to see it as just a proposition with explanatory value. In the case above, it might in fact be right to say this was an explanation in the second sense. The example above does seem to provide *some* explanatory information, even though eating oysters was by no means sufficient cause for my windscreen being smashed. However, when my friend asks me *why* an event happened, he is not looking for any old piece of causally relevant information – only what Lewis would call the most 'salient' bits of information in the circumstances, and bits of information hitherto unknown by my friend.

Not all propositions expressing causes, then, are considered to be good explanations (in the 'act' sense) for their effects – but it seems clear to me that causes are good candidates for explanations. But are all explanations causal explanations? When it comes to explaining events, David Lewis thinks so. For Lewis:

LE *To explain an event is to provide some information about its causal history* (Lewis 1986: 217)

For Lewis, an event's causal history consists of the vast number of perhaps infinitely long causal chains that lead to it, and when we are asked *why* an event happened, we should take it as a request for us to provide a part of that causal history. But this *seems* to rule out a number of intuitively plausible explanations. For example, laws might be thought to explain events. I might cite the law of gravity as an explanation for my wallet falling to the ground, for example. But Lewis can respond that either (a) I have failed to explain this event, as the recipient is already aware of this information (and hence the information is not salient), or (b) if it explains at all, it does so as a part of a causal explanation - 'it provide(s) a peculiar kind of information about the causal history of the explanandum' (Lewis 1986: 222).

Given my conclusions regarding the concept of causation, I can also cash out the causal conceptual analysis of explanation in terms of counterfactuals. For Lewis an event is explained in terms of causal dependence, which he takes 'to be counterfactual dependence, of a suitably non-backtracking sort, between distinct events' (Lewis 1986: 216), where the antecedent explains the consequent. This counterfactual dependence already demonstrates how laws must play a role in explanation, as laws are required in order for us to judge the truth-values of counterfactuals; but this counterfactual account of explanation, I think, also captures why we think 'the law of gravity' can serve as an explanation in itself. A law can, on the face of it, feature as the explanatory antecedent in a counterfactual where the consequent is dependent on the explanands. For example: 'if the law of gravity failed to hold for the few seconds after I let go of my wallet, then my wallet would not have fallen to the ground'.

Whether or not laws are suited to featuring in these counterfactuals from a metaphysical perspective, or whether it is feasible to judge the truth-values of these counterfactuals is another matter, but conceptually speaking at least, to ask 'what would have happened if proposition, *p*, had not been a law?' is a reasonable question, so it seems to me that (conceptually) laws can be explanations in themselves (note that Lewis does not deny this, he just states that a law in an explanatory context must also give us part of the explanandum's causal history).

This discussion of explanation in cause and law will be continued in later chapters, but for the time being I would like to conclude that an adequate act of explaining will provide the recipient with the most salient piece of explanatory information - where, just as in the case of causal dependence, explanatory information can be cashed out in terms of counterfactual dependence. Arguably, all the truths upon which the explanandum depends have explanatory value of some sort (some weaker than others), but many of these pieces of information will be unsuitable for use in acts of explaining.

2.11 The Distinction Between Laws of Nature and Causal Laws, and an Introduction to Regularity Theory

In this thesis I regularly use the terms 'causal law' and 'law of nature', but there is an important distinction between the two as their compatibility with different metaphysical theories of causation varies. This will become particularly apparent in chapter 3, which focuses on regularity theory of cause and law.

By *causal law*, in the broad sense, I mean a proposition or rule which all relevant *object-involving causal processes* abide by. For merely descriptive purposes, it may be helpful to think of these propositions as applying only to object-involving interactions, where one event *seems* in some way to 'produce' the other[24]. For example, it may be a causal law that hydrogen combusts in the presence of oxygen and an open flame, as bringing a lit match to the gas seems to produce the explosion. *Why* this is the case (should there be a reason) is a matter for debate, and depends partly upon which metaphysical view of laws one adopts.

The set of causal laws is a subset of the set of laws of nature, but the latter is not entirely constituted by members of the former. Causal laws, as stated above, are rules applying *specifically to object-involving causal processes*; event *a* causing event *b* would involve a causal law. Laws of nature, need not directly involve causal processes (although there may be causal processes underlying them). It may be a law of nature, for example, that all ravens are black, or that all electrons have charge -1 (note that this cannot be an analytic/*de dicto* truth, it must be true *de re*). This would not be a causal law, as no immediate causal process is involved - neither ravens, electrons, colours nor charge-values are processes. A law of nature can 'be thought of as a universally quantified conditional, $\forall x(Fx \rightarrow Gx)$, which says that anything that has the property F

[24] We will later see that this 'producing' connection is sometimes claimed to be an unjustified assumption, where the 'producing' aspect of the causal interaction is 'in the mind', not in the objects themselves

has the property G.' (Mumford 2007: 42). This may include the diachronic causal laws, but also synchronic, non-causal laws of nature.

The definitions of 'causal law' and 'law of nature' provided above are indicative of a difference between the regularity theory of causation, and the regularity theory of laws of nature. As Armstrong explains (Armstrong 1983:11), those that support the regularity theory of causation must consider a causal connection to involve both the reduction of cause to law (that is, of the universally quantified conditional variety), and the reduction of laws to mere regularities. Reducing causal connections to nomic connections of this kind would mean accepting that causal 'laws are nothing but regularities in the nature of things'[25] (*ibid*).

However, it is possible to dismiss the regularity theory of causation, whilst accepting the regularity theory of laws of nature, by denying that one can reduce cause to law; that is, by claiming that there is more to causation than nomic connection[26]. A regularity theorist of laws of nature could accept $\forall x(Fx \rightarrow Gx)$ to adequately represent laws concerning ravens and blackness and so forth, but deny that the same can be done for causation. The regularity theory of causation thus entails the regularity theory of laws of nature, but the regularity theory of laws of nature does not entail the regularity theory of causation. Indeed, this is evident in David Lewis's work[27].

The list of five conditions for laws set out in section 1.9, and the short discussion of causal laws and laws of nature that follows, is far from complete, but a complete list and explication of the two concepts is not required for the time being. I have provided these only to give some idea of the kind of propositions we will be considering in detail in later

[25] Note that 'nomic' here refers to what the regularity theorist would consider nomic, as opposed to what Armstring would consider nomic. A mere regularity is not a nomic connection for Armstrong.

[26] Of course, even those that deny that causal laws are reducible to laws of nature, would accept that the realisations of many causal laws are consistent with the universally quantified conditionals of the regularity theorist.

[27] Lewis holding a regularity theory of laws, but a counterfactual account of singular causation.

chapters. I believe these propositions will turn out to be central to both the conceptual and metaphysical analyses of causation.

2.12 Conclusions

In this chapter I have considered a number of conceptions of causation: Mackie's conditional account - one that picks out the actual events we consider to be the causes as insufficient but non-redundant part(s) of an unnecessary but sufficient condition of their effects; Lewis's counterfactual account that determines which events are causal by reference to a close possible world, in which that world is left to evolve in accordance with the laws of the actual world from the moment it diverges from actuality; and the dispositional account which states that a causal interaction occurs when dispositional properties are manifested under certain stimulus conditions. I have claimed that each of these accounts is closely linked with laws, and this should come as no surprise when we consider the function of causal statements. Kim suggests the 'common contexts in which we engage in causal talk include:

1) ...*explain(ing)* the occurrence of particular events.
2) ...*predictive* usefulness
3) ...the power to *control* events
4) ...the attribution of *moral responsibility*, and *legal responsibility*; and
5) ...use in special technical senses in physical theory. (see Kim 1973b: 572)

With the possible exception of 4), these functions look to be extremely similar to the functions we would commonly assign to laws.

Importantly, though, it seems to me that neither in our concept of causation, nor in our concept of law, does metaphysical necessitation need to be found.

Part 2

Chapter 3: Hume's Regularity Theory of Laws of Nature

3.1 Introduction

The regularity theory of laws of nature is typically split up into two versions: the 'naïve regularity theory' (Hume) and the 'sophisticated regularity theory (SRT)' (Lewis), otherwise known as the 'best systems analysis'. This chapter is initially concerned primarily with the 'naïve' version of the thesis.

In this chapter I first take a brief look at the desiderata of a metaphysical account of laws of nature, and then explicate the Humean regularity theory of laws in more detail; problems with Humean accounts of laws in general, and criticisms more specific to the regularity theory of laws will then be outlined, and finally I propose some solutions to these problems. In particular I focus on the problem of induction, and argue that this should not be considered any more problematic for the Humean than it is for the necessitarian. I also show why Humean laws, contrary to the necessitarian claims, have significant explanatory value. However, I conclude that there are at least three problems which demonstrate that the naïve regularity theory of laws must be substantially altered if Humeanism is to survive in any form.

3.2 The Desiderata for Laws of Nature

It seems to me that we, as human beings, have a concept of laws of nature as well as of cause and effect; we are able to pick out the kind of statements that can, and those that cannot be law-statements. Accompanying this ability to pick out laws from non-laws come a number of intuitions as to what a law is - the semantic properties law-like propositions can instantiate, the semantic properties no law-like

propositions instantiate, and the role the aspects of reality corresponding to these law statements play in the world. Of course, it might turn out that not all intuitions with respect to laws are metaphysically viable, but to begin any discussion of laws it seems appropriate to take a look into what conditions we would, in general, attribute to laws of nature. In chapter one, following van Fraassen, I made the following suggestions:

1. Given the variety of semantic structures genuine laws can take (functional, non-functional universal generalisations, symmetry principles, propositionally inexpressible mathematical equations) it would be unwise to propose a set of necessary and sufficient conditions;
2. laws *must* hold universally -- omnitemporally and omnispatially;
3. laws must be able to play certain pragmatic roles; in particular, to predict future occurrences, and to explain past occurrences. There is a clear link here with our inductive practices. A good account of laws should certainly help justify many of the inductive inferences we make. Again, as I suggested in chapter one, I believe the explanantia of particular events to be their causes, but nevertheless the laws should help explain the similarities between the causes and their effects;
4. laws must be objective; that is, mind-independent. Laws are out there for us to discover, whether or not we are successful in doing so. Genuine law-statements are those propositions corresponding to this mind-independent aspect of reality;
5. no analytic truths count as laws;
6. the vocabulary used in law statements should include only natural properties.

In this chapter I consider some of these conditions in more detail, and whether or not the regularity theory of laws satisfies them.

3.2 The Regularity Theory of Causation and Laws of Nature

The naïve regularity theory of laws of nature states that ‘‘All Fs are Gs’ is a law iff All Fs are Gs, where the latter is a Humean uniformity’ (Armstrong 1983: 13). As was briefly mentioned in chapter 2, for those who hold the most basic version of the regularity theory of causation there

is no more to causation than a nomic-connection, and furthermore that laws are just regularities: if all events of type-F are followed by events of type-G, then 'all events of type-F are followed by events of type-G' is a law of nature - it is also, if we hold the regularity theory of causation, a causal law, and every instance of this regularity counts as a causal interaction. If placing salt in water always results in the salt dissolving, then it is a law that all salt which is placed in water dissolves. Any instance of salt dissolving when placed in water is thus a causal interaction. This is the view David Hume advocates.

Metaphysical accounts of singular causation thus become redundant when looking for a metaphysical account of causation. What determines a causal interaction for naïve regularity theorists is set by the criteria for causal laws. Note that this is not to say accounts of singular causation would be redundant in a conceptual analysis, even if we do hold a regularity theory of causation. It seems to me that it may well be the case that our concept of causation is focused on singular causation, but that from a metaphysical perspective, any work over and above discovering the best metaphysical account of laws is unnecessary. As we shall see, though, the naïve regularity theory fails to provide a satisfactory account of laws.

The naive regularity theory faces numerous objections. It would of course be possible to just consider only the knock-down objections and to reject the thesis purely on their account. However, it will be useful to look at each of the popular objections in turn, and demonstrate how I believe many of them can be resolved. A fairly exhaustive list of these objections is provided by David Armstrong in *What is a Law of Nature*, so I look at each of these in turn.

3.3.1 The Problem of Induction

One of the desiderata of laws of nature is that they should help support our inductive inferences; that is, we should be able to use laws to predict future occurrences based on past observations. Typically, one should be able to infer a law from previous observations, and then use this law to

predict what will happen were certain circumstances to arise in the future; that is, if we have observed all events of type-F to have been followed by events of type-G in the past, our theory of laws should (in suitable circumstances) allow us to infer a law that will justify our prediction that were an event of type-F to occur in the future, it will be followed by an event of type-G.

Armstrong argues that, given that the Humean sees 'the true form of an inductive inference (to be) simply an inference from the observed cases to the unobserved cases. And, given that the law is just the observed plus the unobserved cases, *that* inference,… is an irrational inference' (Armstrong 1983: 53). The Humean, according to Armstrong, is committed to inductive scepticism[28].

Although this section is primarily concerned with *laws of nature* and not specifically *causal laws*[29], the best indication of the naive regularity theorist's commitment to inductive non-scepticism is found in Hume's discussion of causation. When Hume discusses the formation of ideas from impressions, or our identification of causes and effects, he himself highlights the inductive inferences involved and the problems that come along with it. However, if Hume were to be (as traditional interpretations imply) a strict inductive sceptic, it seems he should conclude that any attempts to identify causes and effects are futile. This would be a strange opinion to attribute to Hume. After all, Hume spends considerable time outlining the conditions under which causes and effects

[28] Armstrong actually attributes this position to Hume himself, but whether or not Hume genuinely believed induction to be irrational is debated. Regardless of the outcome of this debate, this paper is concerned with what the 'Humean', or perhaps rather less ambiguously the 'regularity theorist' should believe, as opposed to what Hume himself believed.

[29] It is possible to dismiss the regularity theory of causation, whilst accepting the regularity theory of laws of nature, by denying that one can reduce cause to law; that is, by claiming that there is more to causation than nomic connection. A regularity theorist of laws of nature could accept $\forall x(Fx \rightarrow Gx)$ to adequately represent laws concerning ravens and blackness and so forth, but deny that the same can be done for causation. The regularity theory of causation thus entails the regularity theory of laws of nature, but the regularity theory of laws of nature does not entail the regularity theory of causation.

should be identified. As Helen Beebee suggests, 'Hume's rules appear to tell us that we *should* seek out hidden causes; but if he is an inductive, and hence causal, sceptic the rules lack any normative force: no purpose is served by acquiring more, or more refined, causal beliefs' (Beebee 2006: 43). It seems to me that if no inductive inferences were in any way justified, no such set of rules would be better than any other.

Whether or not Hume himself was a *non*-sceptic about induction is inconsequential. What is important is that he, as a Humean, *should* have been - if a regularity theorist wishes to identify any laws of nature (even with an admission of fallibility), he *must* be a non-sceptic about induction.

3.3.2 Why a Regularity Theorist Must Reason Inductively to Identify Laws of Nature

A regularity theorist believes that when we identify a law of nature, we do so by observing a constant conjunction between certain properties (the property of being a raven, and the property of being black, for example). The observed instances are, of course, constrained to our present and past experiences, but according to the regularity theorist, the constant conjunctions that make up a law of nature must hold omnitemporally and omnispatially.

It follows that if the regularity theorist is to justify his identification of laws, he must justify his belief that the constant conjunctions identified will hold across all spatio-temporal regions. He is making conclusions about the entirely unobserved future, from the partially observed past, and in so doing is committing himself to the rationality of inductive reasoning.

It may be argued that the regularity theorist need not *identify* laws of nature in order to maintain his primary beliefs about what a law *is*. For a universal regularity, *l*, to hold; that is, for there to be law, *l*, nobody needs to actually know, or even believe *l* to be a law at all. Even when a law *is* identified, the regularity theorist must accept he may be wrong. The regularity he thought was a law can always turn out not to be, as the

regularity can (as far as he knows) always break down at some point in the future, or may even have already broken down at some unobserved point in the past. So whether or not a regularity theorist identifies a regularity as a law has no bearing on whether it actually *is* a law. Nevertheless, it is clear that the regularity theorist attempts to identify laws, so the problem remains.

3.3.3 Is the Humean Defeated by Induction?[30]

The 'problem of induction arguments' raised by necessitarians are said to apply not only to the regularity theory, but to all Humean theories. Let us take the Humean view of laws to be the view that laws of nature are either regularities, or regularities described by universal quantifications that are part of a best system of law-statements. As I have said, like anyone, a Humean in the world has to rely on inductive inference to arrive at beliefs about which particular natural laws there are. Armstrong argues that Humeans have a special problem with induction: given their conception of law, induction is an irrational inference. I look at Armstrong's metaphysic in more detail in chapter 5, but for now it will suffice to say that for Armstrong, a necessary condition for the rationality of inductive inference is that one accepts his conception of natural laws; that is, that laws are higher-order facts linking first-order universals by his 'natural necessitation relation', or the N-relation. In other words, the metaphysical glue linking the universals, say, F and G, constitutes the non-accidentality of the regularity, that all Fs are Gs. Armstrong does not claim to solve the problem of induction as such - he just claims to show that a necessary part of any justification will be his conception of laws. So Humeans miss out on justifying induction since for them, there is no metaphysical glue underlying the regularities of the spacetime continuum; that is, the Humean mosaic. Metaphysically speaking, there is just the mosaic.

In this section I argue that Armstrong's attack on Humeanism fails. Armstrong has not shown that one needs higher-order necessitation

[30] Section 3.3.3 through to section 3.4 are forthcoming in Smart, B. 'Is the Humean Defeated by Induction' *Philosophical Studies*

relations to validate induction. There are two lines of argument that correspond to two issues in relation to induction: the old problem and the new problem. The first is that Humeans will have a specific problem with Goodman's new problem of 'grue'. I show that here Armstrong is simply confused, and indeed, liable to fall to his own objection. The second argument is really concerned with the old problem of induction. The central claim is that universally quantified statements about regularities, such as *All ravens are black*, have no explanatory value. Armstrong's argument here is marred by confusion about what we are meant to be explaining. When we get clear about what we are explaining, (in this case we are explaining facts of observation and not the colours of particular ravens), we see that such statements about regularities can have explanatory value. Furthermore we see that Armstrong himself must admit that they do if he is to allow for there to be explanations for many of the every-day events we are asked to explain. A second feature of the second argument concerns chains of explanation. I claim that the chain of explanation need not end with some strange entity like the necessitation relation, but can end with a more comprehensible regularity. I conclude that the regularity theorist's alternative to N(F,G), which I term the regularity relation between universals, R(F,G), is the best explanation of our worldly facts, and that this provides the regularity theorist with *at least* as much right to reason inductively as the Armstrongian.

None of this shows that we have solved the problem of induction for the Humean. It just means it hasn't been demonstrated that they cannot. I end this section with a not-implausible *a priori* attempt at providing a justification of induction in Humean terms. This attempt is grounded on the probabilistic justifications provided by Williams and Stove, but sadly, by current standards, fails due to its supposedly unjustified assumption of the proportionality syllogism – however, further work on probabilistic justifications of induction might prove my attempt successful after all.

I said that Armstrong has two lines of objection: one for the old problem, one for the new. I begin with the new.

3.3.4 Armstrong on the New Problem

When I discuss the new problem of induction, I refer to the problem raised by Nelson Goodman's (Goodman 1979: chIII) 'unnatural' predicates. Goodman asks us initially to assume that if all emeralds observed before the year 3000 have been green, it is rational to conclude that all emeralds are green. He then introduces a new predicate, 'grue' – an object is grue if it is first observed before the year 3000 and green, and blue if first observed thereafter. All observed emeralds have been green, but they have also been grue, so why not conclude that all emeralds are grue? Of course then we'd have to predict that all emeralds first observed after the year 3000 will be blue, and that would contradict the original projections. The grue problem is one that applies to the rationality of induction in general, and Goodman (see Goodman 1979: ch. 3) provided arguments to suggest we can solve the problem by accepting only natural predicates like green and black in law-statements – which can be identified because they are the best entrenched; that is, predicates like green and black are routinely used by the general populus, and 'unnatural' predicates like grue are not. Restricting the Humean's inferences to this kind of predicate would solve the problem, but Armstrong argues that it is 'impossible to see how the new principle [of restricting inferences to natural predicates] is to be justified' (Armstrong 1983:58). Although he says nothing more as to why it couldn't be justified, I assume the reasoning goes something like this: Whereas Armstrong may rule out grue-like predicates as only natural predicates are, in Goodman's words, 'well-behaved predicates admissible in lawlike hypotheses' (Goodman 1979:79), the Humean does not hold a law-like hypothesis, so this kind of response is unavailable to him. However, I believe the Humean, too, can

make the claim that grue-like predicates cannot be used in law-statements[31].

It is true that the emeralds in our sample have been both green and grue, and so the inferences to 'all emeralds are green' and 'all emeralds are grue' look to be equally well supported, but why can't the Humean also appeal to the natural/unnatural distinction? One of the most defended criteria for being a law-statement (for Humeans and necessitarians alike) is that it only includes natural predicates, so if any proponent of a view of laws is to have a plausible metaphysics of laws, he must be able to give an account of this distinction. Armstrong appeals to objective similarities in the form of universals, but it's not entirely clear why universals are required to pick out similarities. David Lewis (Lewis 1986: pp59-61) suggests that naturalness could also be seen either in terms of objective structural similarities, or simply just primitive naturalness, and makes no reference to universals whatsoever. *All* metaphysical theories are left with primitives, and it seems to me that there is no reason why a Humean cannot take the natural/unnatural distinction to be a primitive fact. For the purposes of this paper I will assume this primitive distinction between natural and unnatural predicates can be made, and thus the grue problem overcome.

It may be, however, that there is an implicit idea in Armstrong that naturalness needs to be connected to law. When we encounter the dispositional essentialist metaphysics in later chapters, we will see that for them, being a natural property is constituted by lawful relations to other properties (Bird 2007). Armstrong's view is distinct from this viewpoint, since for him, naturalness is determined by something that is not inherently connected to law. His view is quidditistic, so for him, naturalness is entirely primitive. In this respect, he is in the same boat as the Humeans in that the identity of a property is not fixed by its causal/nomological role. For Armstrong, then, natural properties cannot be picked out by their connection to laws.

[31] Or any other unnatural predicate for that matter.

3.3.4 Regularities and Explanation

I now move to the second objection, which is firmly couched in terms of the old problem of induction. Armstrong argues that the rationality of induction is a necessary truth, not just analytically, but for some 'deeper reason' (Armstrong 1983: 54). He proposes that the necessitarian can rationally predict the continued uniformity of nature by inferring the 'natural necessitation relation between universals', N(F,G), to be the best explanation for our observations, and that this relation justifies our inductive inferences. According to Armstrong, however, there is no way for the Humean to justify inductively derived predictions about the future. He claims their predictions about unobserved events are not grounded by inference to the best explanation (IBE), but based solely on the 'pattern of inference: observed instances to unobserved instances'; where *e* (the observed instances) is inductive evidence for *h* (claims about unobserved instances). Armstrong suggests that the Humean reasons as follows:

1. $e \rightarrow (e + h)$

2. $(e + h) \rightarrow h$

3. $e \rightarrow h$ [From 1, 2, by transitivity.]

The inference from *e* to *h* is, according to Armstrong, to be regarded as an irrational inference (although I will later claim that (a) this pattern is common to *all* inductive inferences (which in my opinion can scarcely be questioned!), and so to say this pattern is irrational is to say that induction is always irrational, since it takes us from the observed to the unoberved, and (b) Armstrong is committed to the rationality of the pattern himself.

Ultimately, though, Armstrong's argument rests on the claim that Humean law-statements - in the form of universally quantified statements- have no explanatory value whatsoever, largely because they do not provide (what I call) the 'metaphysical glue' required to keep nature

uniform. The focus of this section is to demonstrate why these conclusions are false.

For Armstrong's pattern of inference to justify *his* inductive inferences, he needs IBE to be a justified mode of reasoning. I have no objection to this, but first we need to look at what this principle amounts to, and how it should be applied in the particular cases I'll be looking at. Peter Lipton suggests that when one is considering which is the best explanation for a phenomenon, one has to consider both which is the likeliest explanation, and 'which would, if correct, be the most explanatory or provide the most understanding'. (Lipton 1991: 59) We therefore need to establish, first and foremost, what phenomena we're actually explaining. When Armstrong asks us to explain why "all ravens are black", he cannot be looking for an explanation for *all* ravens (omnitemporally and omnispatially) being black, largely because this information is not available to him prior to the inference of a natural necessitation relation holding between the universals of ravenhood and blackness. *That* all ravens are black is just an implication of his explanation.

If we required a reason for all ravens being black then perhaps Armstrong might have a better chance of explaining this phenomenon than the Humean (although this is not obviously true), but a Humean *doesn't want* there to be an explanation for this over and above regularities, as for him the universal blackness of ravens is just determined by the distribution of local particular matters of fact. What requires explanation here is why all the ravens in *our sample* are black, and one must appreciate that a perfectly good explanation for this might not serve as an explanation for why ravens 'in general' are black.

Not only do I contend that pure universally quantified statements concerning regularities can serve as good explanations for our observations, but also that they are frequently the best explanation available. Furthermore, it turns out that Armstrong's *own* pattern of inference takes the form of $e \rightarrow h$, which he himself regards as irrational. I

will thus argue that his argument thus fails on three counts: firstly that to allow for widely accepted explanations to have explanatory value he requires universally quantified conditionals to have explanatory value, secondly that as his pattern of inference stands, Armstrong's account fails by his own standards, and thirdly because (even if we could make sense of the N-relation) Armstrong's explanations are not the *best* explanations (he needs them to be if he is to appeal to IBE).

If Armstrong's metaphysical view of laws is right, then $\forall x(Fx \rightarrow Gx)$ is true at every world in which N(F,G) holds. The strange metaphysical glue that is the natural necessitation relation sticks the instances of universals together, such that wherever something instantiates the property F, it also instantiates the property G. N(F,G) thus entails the above universally quantified statement. If we consider the universals 'ravenhood' and 'blackness' to be joined by the natural necessitation relation, N(R,B), we know that instances of ravens will always be black. But, as we shall see, not all explanations require this metaphysical-glue aspect – regularities alone can often do the job.

That a regularity can explain its instances should not be particularly counter-intuitive. After all, we use regularities as predictive and explanatory tools all the time. Take the following example: Political polls tend to provide us with accurate predictions of election results, and it seems perfectly reasonable to explain the results of the polls by appealing to the proportions in the total population of voters. We *could* look for a further reason for the results of the polls; perhaps one political party had particularly unpopular policies, but (a) the results of the polls simpliciter are sufficient for us to predict the election results, and (b) the proportions in the total population do seem to explain the results of the polls. Do the proportions explain *why* people choose to vote the way they do? No. But this is not what's required when explaining why our polls provide accurate predictions of election results.

Whereas Armstrong seems to think an explanation must be metaphysically 'meaty' - that the explanans must involve a kind of metaphysical glue ensuring the universals are tied together (and that his natural necessitation relation provides us with these kind of explanations), there are thousands of examples of proportions in a population being used to explain the proportions in a sample. If an account of what it is to be an explanation denies us the ability to explain proportions in samples by appealing to the proportions in the relevant population, then that account is clearly at odds with our concept of explanation.

We can also see this when we look at explanation from the bottom up, to see what kinds of truths need explaining (as opposed to what kinds of things can explain) - here we find many explananda that Armstrong cannot deal with, as the explananda are not captured by relations between fundamental properties. Consider the following examples: "Only 20% of the people in my street are from ethnic minorities." "Girls at Victoria College performed better than boys in their finals." Presumably we can ask for explanations here, but given that there are no fundamental universals for the natural necessitation relations to hold between, Armstrong cannot provide them (and note that he cannot appeal to a causal explanation, as, for Armstrong, causes are just instances of laws). The Humean, on the other hand, can respond with regularities: "actually, only 20% of people in Britain are from ethnic minorities, so it should not be surprising that only 20% of people in your street are"; "most female students are more diligent than boys", and so on.

Finally, Armstrong's initial objection was that the Humean's pattern of inference was reducible to the 'irrational' $e \rightarrow h$ - but if we look at Armstrong's own pattern of inference we can clearly see that he himself is committed to this pattern, as ultimately Armstrong's inference is an inductive inference, and *all* inductive inferences take this form. Inference to the Best Explanation, although a special kind of inductive inference, is an inductive inference nonetheless. It is not *deductively* derivable from a formalism of the observed instances, and so it makes the jump from the

observed to the unobserved. In this case, the inference is not to the observed instances plus the unobserved instances (*e+h*), but from the observed instances to the (arguably) *unobservable* natural necessitation relation between the respective universals. If the Humean's 'observed to unobserved' pattern of inference is irrational, then so is Armstrong's[32].

So far I have mainly focused on how the Humean can explain our *observed samples*; that is, the colour of the ravens we have observed, but we might well ask for a reason why *all* ravens are black - why ∀x(Rx→Bx), not just why 'all ravens in our observed sample are black'. In the next section I will endorse a hierarchy of explanations, all of which are regularities explaining regularities lower down in the hierarchy. By the very nature of this hierarchy, there has to be a highest-order regularity, and hence a brute, unexplained explainer. This is often presented as an objection to the Humean, but I show that the Armstrongian explanatory chain also ends with a brute fact - namely the natural necessitation relation. I claim the explanatory chain *should* end with a regularity, and not this mysterious relation between universals.

3.3.5 Higher-Order Regularities as Explanations

Nicholas Everitt (Everitt 1991: 206-208) pointed out that nearly all regularities (but not all or else we get a regress) can be explained in terms of higher-order regularities. For example the law of nature 'all water has mass' can be explained by the higher-order regularity: 'all molecular substances have mass'. Necessitarians are likely to claim, however, that explaining a regularity in terms of another regularity is not sufficient, for when a regularity is reached for which there is no higher-order regularity there are no further explanations.

Intuitively, saying a particular cricket is green because 'all crickets are green' (if this were the case) is not a sufficient explanation for the

[32] Let me emphasise that this is not to say that I agree with Armstrong in thinking *e→h* is always an irrational pattern of inference. I believe that inductive reasoning *is*, when performed in the right way, perfectly rational. The conclusion *h* can follow from *e*, not by deductive logic, but by whatever inductive logic will turn out to be.

greenness of crickets. It would be far more helpful to appeal to the evolutionary process, whereby green crickets survive *in virtue* of the camouflage greenness provides them with. A higher order regularity like 'all non-camouflaged insects get eaten by birds' may be proposed by the regularity theorist, but once again, no real explanation is given for why this is the case. It seems that regularities in themselves do not have the explanatory power required by the necessitarian as they must always end in a brute fact, so explaining regularities in terms of other regularities is, at least at the top end of the scale, a futile exercise.

Everitt concedes that his appeal to higher-order regularities may be criticised in this way, but he continues his attack on the scientific approach by claiming the necessitarian *also* runs out of explanations at the top end of the scale. A useful analogy can be found in Hume where he makes a similar claim with respect to causation. He wrote that in following a causal chain of events (backwards), 'Every link of the chain wou'd in that case hang upon another; but there would not be any thing fixed to one end of it, capable of sustaining the whole; and consequently there would be no belief nor evidence.' (*Treatise*: Bk 1, IV) If I take a causal chain of events, A-B-C-D, whereby a subsequent event can be explained by the antecedent event, event D can be explained by appeal to event C. But what happens when we get back to event A? The physicist would normally associate event A with the Big Bang, but how can we explain the coming about of the Big Bang? Everitt claims, I believe quite rightly, that some 'version of an ontological argument from essence to existence' (Everitt 1991: 208) would have to be accepted.

The analogy above focuses on causal explanation, but the same applies with Armstrong's nomic explanations. For Armstrong, N(F,G) is supposed to explain why this F is a G, and allows us to predict that all future Fs will be Gs, but what explains this fact? For Armstrong, the natural necessitation relation, N, is a universal just like F and G, and so just like F and G it should be able to change its causal role across possible worlds; that is, in another possible world N plays the F-role and F the N-

role. Alexander Bird shows that any attempt to solve this problem results in a vicious regress. Bird states that if R(F,G) denotes a regularity between Fs and Gs, then 'N(F,G) necessitates R(F,G) (for all F,G). We may symbolise this N'(N,R)... N(F,G) cannot explain R(F,G) if it is just a regularity that whenever N then also R' (Bird 2005: 151) , so we introduce this further (third order) necessitation relation, N', to explain this regularity. But of course then we need an explanation for why whenever N'(F,G) then also N(F,G), when it looks like this must also be a regularity, so we introduce N'', and so on. Ultimately, if Armstrong wants to use laws of the form N(F,G) as his explanation for Fs being Gs, he may as well take the initial N-relation between first order universals to be the primitive explanans (the function of which, I should add, we have no idea about). It seems to me that if Armstrong must concede that his explanatory chain finishes with an unexplained fact, he cannot just appeal to this 'explanatory chain finishes' objection to stamp his authority. There is thus no obvious undermining of the Humean's 'in principle' ability to explain the validity of induction.

3.3.6 An Attempt by the Humean Foiled

The Williams-Stove (Williams 1947, Stove 1986) argument from The Law of Large Numbers (LLN) demonstrably shows that for any finite population, of all the logically possible large samples in that population, most of those samples will have proportions that resemble the proportions in the total population - so given any random large sample[33], it is statistically likely that the population from which that sample was taken will have proportions that resemble the proportions in the sample. It follows that future samples are likely to resemble past samples.

LLN tells us that if we choose 3000 ravens at random from a population of a million ravens, half of which are black, half white[34], the

[33] Importantly, 'random' in this sense must mean that every one of these logically possible samples must have an equal probability of being drawn.

[34] Note that a 50/50 proportion will give the lowest probability of a representative sample. A population of 100% ravens will of course give a 100% chance of a representative sample

statistical probability of the proportions of that sample being within 3% of the proportions of the total population (between 47% black and 53% black) is greater than 0.9 (Stove 1986: 70). Given there is a statistical probability of more than 0.9 of any one sample being representative of the population, Stove reasons that we are justified in projecting the same proportions found in the sample we draw on to the population as a whole - this supposedly provides the probabilistic justification for inductive inferences. Given that we need epistemic justification for induction, there is an implicit assumption here that, given the statistical probability of drawing a 'representative of population' sample-type is greater than 0.9, the *epistemic* probability of that sample being representative of the population is *equally* high - Stove is assuming the proportionality syllogism[35].

The calculation of the statistical probabilities is a very simple *a priori* matter. One merely calculates the number of logically possible 3000-fold samples in the population, and then the number of these possible samples whose proportions of black-to-white ravens fall within 3% of the proportions of the total population (I'll call these samples 'representative samples'). We then divide the latter by the former to get our probability. It is just a mathematical truth that more of these samples are representative than non-representative[36], and so for any random large sample of a finite population, one is likely to get proportions closely resembling the proportions of the total population. It is important to note that once we have a sample greater than 3000, the size of the total population does not have a significant effect on the proportion of 3000-fold samples representative of the population; that is, if we have a total population of a hundred trillion instead of a million, the majority of samples will still be representative. This may sound counter-intuitive, but

[35] This tells us that if we know the number of Xs that are Ys in a population of Xs (where we know the size of the population), then our epistemic probability of a random X being a Y is the number of Xs that are Ys divided by the population size. If 60% of professors are men, our degree of belief that an unknown professor (of whom we have no further information) will be a man should be 0.6.

[36] Assuming an equal probability of choosing any possible sample.

it is easily shown to be true mathematically[37]. It follows (again, if we adopt the proportionality syllogism) that we can justify many inductive inferences. But can the Humean use this principle?

Those who uphold Humean Supervenience (which will be discussed in greater detail in chapter 4) believe 'that all else supervenes on the spatiotemporal arrangement of local qualities throughout all of history, past present and future (Lewis 1994: 474).' For Humeans like Lewis, all contingent truths are true in virtue of these patterns of fundamental property instantiations and the fundamental relations between these instantiations, and in virtue of these alone. For the Humeans there are no natural necessitation relations, nor any other kind of metaphysical glue that ensures the uniformity of nature. But although our observations are just patterns of property instantiation in a restricted spatiotemporal region, these observations are nevertheless a sample of the population as a whole. If the LLN works as a means of justifying inductive inferences, it seems to me that we can easily apply this within a Humean framework as the LLN is a purely statistical device, and requires no commitment to any kind of necessary connections.

Unfortunately for the Humean, the Williams-Stove argument faces numerous objections that cannot (or at least have not) be overcome, and so the aim of finding an indubitable justification of inductive reasoning cannot be met without first dealing with the Williams-Stove argument's problems. I will first outline these objections, and then move on to the more pressing question of whether or not Armstrong has more right to reason inductively than the regularity theorist.

Firstly it is argued that a probabilistic justification of induction requires the *epistemic* probabilities to be high, but the LLN can only demonstrably show the statistical probabilities. As I have already mentioned, the Williams-Stove method makes the non-trivial assumption that the statistical and epistemic probabilities will be the same, but Bruno

[37] It has been argued that infinite populations are problematic, but these will not be addressed as there are far stronger objections than this.

de Finetti (1964) argues that (at least) epistemic probability should be equated with strength of belief, where our degree of belief is judged by 'how much we are willing to bet' for the prize of $1[38]. The strength of our belief (and thus the epistemic probability of our inferences), though, is in force *prior* to the calculation of statistical probabilities, and of course this prior degree of belief is already (at least partly) determined by our innate assumption that Nature is uniform.

Secondly, the argument requires that the probabilities of drawing different samples are 'exchangeable'. The probabilities are 'exchangeable' only if for any two samples with the same proportions, the same degree of belief that they would match those proportions applies to both samples. But this is usually not the case with the kind of inductive inferences we make, largely because causal conditions often change during the course of our sampling: if the weather changes from dry and sunny to wet and windy during the course of a golf tournament, my degree of belief in an unknown individual beating his handicap will depend on when he played. Suppose two golfers, Ben and Jerry, both broke their handicap by 2 shots. Ben played when it was sunny, so my prior degree of belief was 0.6, but Jerry played when it was windy and raining, so my prior degree of belief was just 0.2. In cases like these the probabilities are not exchangeable, and so the Williams-Stove method cannot be applicable.

Thirdly, the LLN requires our sample to be a random one, but it is not at all obvious that we can take any observations to be random as we are restricted to our locations in space and time. Drawing from Giaquinto (Giaquinto 1986 p.614), imagine that the only ravens we have observed have (a) been black, and (b) located in a single valley. We have also observed many other species of animals (some in this valley, but some in others), many of which have differed in their colour from one environment to another. Given we know from experience that the colour of animals may (at least partly) vary from one habitat to another, we would probably

[38] Subsequent subjectivists have substituted, for this monetary assessment, bets of utility, on the grounds that the value of money differs from the rich to the poor.

not be willing to bet large sums of money on ‘all (or most) ravens are black’ being a true statement. Now of course this is an example where we already know our sample is not random, and so one might object that we can assume our sample *is* random unless we already know otherwise. But sadly no observational samples are truly random, as what we can observe is always restricted by our spatial and temporal location. Furthermore, even if we knew that 95% of samples are representative, what justifies us in assuming *our* sample is one of the representative ones? Perhaps this can just be taken as a primitive fact – what justifies us in thinking ours is a representative sample is simply that most samples are (see Campbell and Franklin 2004). Nonetheless, I think the randomness objection should be taken seriously.

And fourthly, when we consider what it is to be an inductive inference, we hope that as evidence accumulates our inferences converge on the truth (importantly though, we should never know when we have obtained the truth of our inference.) According to Kelly and Schulte inductive reasoning has similarities with algorithmic computation, except the latter confers certainty. They write that ‘This certainty derives from two factors (1) a logical guarantee that the algorithm will produce the right answer on each input in a specified class, and (2) the fact that the algorithm halts, thereby signalling to the user in an unambiguous way what its output is’ (Kelly and Schulte 1995: ‘The Problem of Induction’ *Stanford Encyclopedia of Philosophy*). The LLN approach accommodates the uncertainty of inductive inferences by rejecting (1) in favour of the probabilistic approach, but this does not help us with our desire for inductive inferences to converge on the truth. Articulating Kelly’s position, Vickers writes:

> Any good account of scientific reason… must classify and account for the complexity and difficulty of inductive inference… [Probabilistic] methods… are incapable of accounting for complexity and the interplay of conjecture and refutation - logical omniscience runs roughshod over the critical distinctions. The probability P($h|e$) of a given hypothesis h conditional on changing evidence e may fluctuate from close to zero to close to one as e accumulates.

> This violates the central principle of convergence: A conjecture if false will be rejected at some stage, and if true will never be rejected. (Vickers, J 'The Problem of Induction' *The Stanford Encyclopedia of Philosophy*)

Kelly and Schulte provide their own account of how we should deal with the problem of induction, but for the purposes of this paper it is enough to show that probabilistic attempts to justify inductive inferences are insufficient, as they fail to respect the principle of convergence.

I have provided four major objections to the Williams-Stove response to the problem of induction, which have not, as far as I can see, been adequately answered by the Humean. But it is important to note that although this probabilistic attempt to justify the regularity theorist's ability to reason inductively has failed, the failure is not one that can be overcome by introducing necessary connections – the objections apply to *all* probabilistic justifications.

It turns out that my probabilistic argument to justify the Humean's right to reason inductively falls short, but that's not to say that future attempts at probabilistic justifications, or work on the proportionality syllogism, cannot prove my attempt, or a variation on it, successful after all. This, however, is work for another day.

In the following sections I examine Armstrong's objections in more detail, and conclude that these do not help motivate his governing laws.

3.3.7 Regularities and Induction

Necessitarians in general object to the Humean's inductive method by stating that the explanans of observed instances cannot include the explanandum; that is, they claim that $\forall x(Fx \rightarrow Gx)$ cannot explain why all the observed instances of Fs are Gs, because $\forall x(Fx \rightarrow Gx)$ is equivalent to 'all the observed Fs and all the unobserved Fs are Gs'. Although these are truth-conditionally equivalent, I have hopefully shown that they are not explanatorily equivalent.

Armstrong claims a natural necessitation relation between universals can be found via IBE to explain our observations, and that given N(F,G) is in some way distinct from unobserved instances, we can use N(F,G) to make inductive inferences about future Fs. But the Humean can use the same methodology to arrive at R(F,G), where R signifies a regularity between the properties F and G - when R(F,G) holds between 'ravenhood' and 'blackness' it can be said that "all ravens are black" (As a matter of fact, the regularity theorist doesn't need to be so bold as to claim that *all* ravens are black to justify his inductive inferences and explain his observations, but 'all ravens are black' would, I imagine, be considered a better explanation than 'all or most ravens are black', and so the rules of IBE require us to stick to the universal generalisation.). Of course if we are justified in inferring that all ravens are black, then we are justified in making projections about the colour of future ravens.

R(F,G), contrary to the more 'spooky' natural necessitation relation, (Armstrong himself writes 'The inexplicability of necessitation just has to be accepted. Necessitation, the way that one Form (universal) brings another along with it as Plato puts it in the *Phaedo*, is a primitive, or near primitive, which we are forced to postulate. (Armstrong 1983: 92)) is easy to comprehend. It is a contingent fact about all Fs which must be inferred through the regular observation of Fs being Gs. Inferring R(F,G) simply requires it to be the best explanation for our observations of Fs being Gs, and as I have shown, this is often the case. If N(F,G) can stand between observation and conclusions about the unobserved for Armstrong (once it has been inferred through IBE), R(F,G) can stand between observation and conclusions about the unobserved for the regularity theorist. The pattern of inference becomes observation→R(F,G), then R(F,G)→projections about future instances, and this, it seems to me, is a far more rational pattern of inference than Armstrong's. It turns out that if Armstrong's mode of reasoning via IBE to justify our inferences is acceptable, neither the necessitarian nor the regularity theorist need appeal to anything more than IBE and R(F,G) to justify their inductive inferences.

In this section I have attempted a probabilistic argument to justify the regularity theorist's right to reason inductively; unfortunately, due to the general failure of probabilistic justifications of induction, this attempt was ultimately unsuccessful. However a complete justification of induction for the regularity theorist is not required to dampen the necessitarian arguments against the Humean. What I have shown should at least deflate Armstrong's objections. I have clearly demonstrated that Armstrong's position fails by his own standards; that regularities can and often *do* serve as perfectly good explanations; and that although it is important to note that I have not solved the problem of induction for the Humean, I *have* shown that the Humean has *at least* as much right to reason inductively as many necessitarians.

Although I believe the necessitarian arguments pertaining to explanation and the problem of induction can be dealt with by the Humean fairly adequately, there are more serious objections to Hume's metaphysic which I do not think can be overcome by the naïve regularity theorist. I outline these below with a view to showing how they can be dealt with by the sophisticated Humean.

3.4 The Problem of Single-Case Uniformities

For Hume (in respect of causal inference), observations of single-case uniformities would never bring about nomological ideas, but we cannot ignore single-case uniformities when approaching laws from the regularity theorist's metaphysical perspective.

When considering laws of nature in connection with observations of single-case uniformities, whereby for any given particular 'x' there has only ever been one instance of 'x' instantiating the property F, and that particular 'x' also instantiated the property G, the proposition $\forall x(Fx \rightarrow Gx)$ is true. However, this would bring about an enormous number of counterintuitive laws. It would be a law of nature, for example, if there had only ever been one person named William Shakespeare (the playwright), that 'all persons named William Shakespeare are

playwrights'. Laws of this kind would certainly not usually be thought of as laws of nature, but without further development, the proponents of the naive regularity theory of laws of nature cannot deny their law-status.

3.5.1 Non-Existent Subjects and Functional Laws

A further problem is that the proposition ∀x(Fx→Gx) is *always* true if no particulars ever instantiate the property F, for this proposition effectively translates as 'either that thing is not an F, or if that thing is an F it is also a G' (Armstrong 1983: 20). Of course if there are no Fs then the proposition is always true, potentially providing us with contradictory laws of nature. For example, the proposition 'All unicorns are 2 meters tall' and 'All unicorns are 3 meters tall' would both be laws of nature.

This may lead us to add the condition that laws of nature should only apply to existent subjects; that is, a law of nature only takes the form ∀x(Fx→Gx) if at some point in time there exists an instance of F. However, the extension of statements of uniformity to require existential quantification if they are to be laws would arguably have the unwanted knock on effect on the many functional laws we would like to include as laws of nature.

3.5.2. Functional Laws

Functional laws of the form X=F(Y), where X and Y stand for variable properties in a certain functional relationship, are common in modern science, but it is not clear that regularity theorists cannot accommodate them. If laws under the regularity theory must be formed 'instantially', then it seems in order to construe a functional law, all possible instances of the variables must be realised.

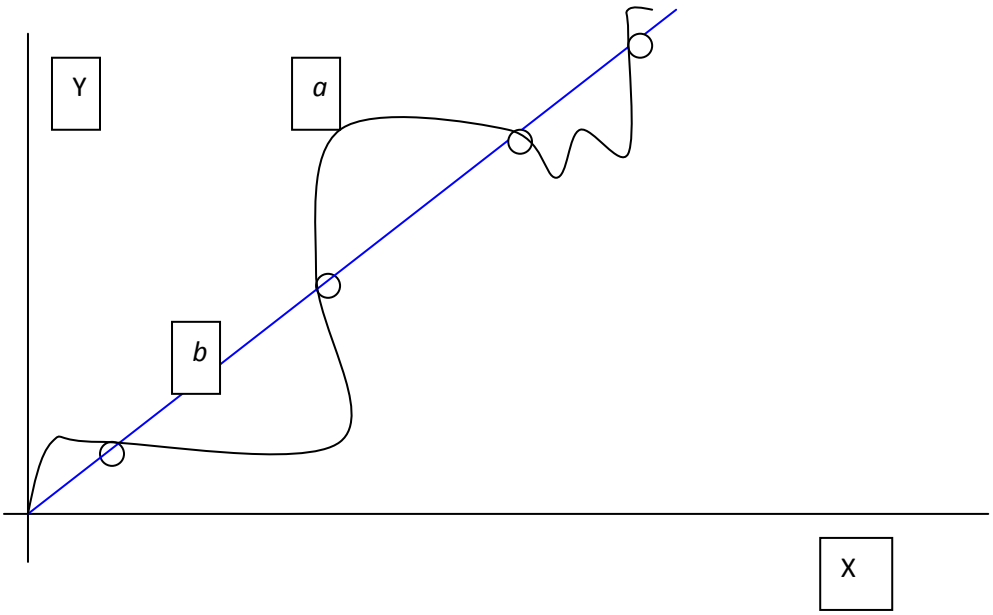
The law F=Ma is a functional law quantifying the relationship between mass, force and acceleration. The force being exerted on an object is directly proportional both to the mass of an object and to the rate at which that object is accelerating. By using the law F=Ma, we can calculate the force required to accelerate an object of one billion kilograms

at 1 m/s^2 to be one billion Newtons. But suppose that no object has ever had the mass of exactly a billion kilograms. Does this mean (under the naive regularity view) that we cannot justifiably calculate the force required to accelerate this object?

Functional laws provide no indication of what instances have been realised, but why should they need to? If the proposition 'all ravens are black' were a genuine law of nature, it would show an omnitemporal/omnispatial relationship between instances of properties within a system; namely ravens and blackness. The law 'all ravens are black' provides no indication of which instances satisfy the law, only that no instances of ravens violate the law. For the Humean, the identification of the law of nature 'all ravens are black' can come from nothing more than the observation of many ravens and inductive reasoning, and this seems to apply to functional laws as much as it does to non-functional ones. We observe the relationship between the two variables in question over time, and see that every instance observed conforms to a certain pattern, quantifiable into a specific functional relationship. Although the values for (unrealised cases of) Y may not be demonstrably *entailed* by X, this value can be inferred by inductive reasoning. The regularity theorist just argues that the counterfactuals are supported by the functional law, and what justifies the inference to the functional law is that mysterious logic which holds between an inductive argument's premises and its conclusion (see Suchting: 1968). We can, on this basis, infer a functional law – the question becomes, what function do we choose?

Unfortunately for the naive regularity theorist, for any given relationship between two variables there is an infinite number of equally viable functional laws. With the law X=F(Y), F (the function) could be inferred by drawing a line on a graph (that plots the corresponding values of the variables) passing through all the instantiated values. We would naturally draw a straight line (see *b*) in the case of direct proportionality, but why should we do this? Drawing erratic, 'squiggly' lines, that also pass through all the instantiated values (see *a*), would represent functional

relationship equally consistent with our data. How could we possibly know which functional relationship is the right one to hold, other than by mere intuition?



Perhaps I could claim, as a naive regularity theorist, that *both* functional laws are in fact real laws in nature, corresponding to the same set of events. If I do this, however, I definitely can't make the counterfactual claims I might want to. I cannot say that "if there were a object of mass one billion kilograms (when there had been no instantiations of an object of this mass being accelerated at 1 m/s^2) it would require a force of one billion Newtons to act on it to accelerate that mass at 1 $m/s^{2"}$, as I would have no justification for applying the F=Ma functional law over any other functional law consistent with the already instantiated instances. Claiming them both to be laws is not a viable option if I want either to be useful.

As David Armstrong puts it:

> The difficulty for the naïve regularity theory is this. It seems natural to think that, although many possible functions are compatible with the data, there is in fact just one function which constitutes the law which actually governs the situation. But the relevant set of Humean uniformities do not logically determine what that function is. (Armstrong 1983: 38)

If we can't choose one function over another, then, the regularity theorists' attempt to support counterfactuals using functional laws will fail. It seems to me, though, that although the kind of logical entailment Armstrong is looking for is absent, it is not unreasonable to think that some functions passing through the actualised instances of a functional law are primitively simpler than others. The straight line is, it seems to me, primitively simpler than the squiggly line. Primitive objective simplicity and inductive logic may give the Humean some confidence in their right to infer the right functional laws. Supporting counterfactuals, however, remains an issue for the naïve regularity theorist, even if we're happy to accept this objective primitive simplicity answer to the functional laws problem.

3.6 Supporting Counterfactuals

One of the most important criteria for laws of nature is that they can be used to support subjunctive conditionals/counterfactuals. I don't want to be able to say[39]just that "all ravens are black explains why this raven was observed to be black", I also want to be able to say that "were I a raven, I would have been black". Critics of the naive regularity theory of laws of nature claim that regularity theory laws are unable to support these counterfactual claims.

3.6.1 Counterfactuals with no Instances

Michael Tooley (Tooley 1987: 50), in a paper discussing reductionism with respect to laws, asks us to consider a psychophysical law connecting neurophysiological states to phenomenological states incapable of being derived from any other laws. The law relates to a certain brain state giving rise to an experience of a particular shade of purple, and applies only to sentient beings on Earth.

Let us assume that in world, *w*, that specific shade of purple is observed at time *t.* What happens when, in another possible world, *w**, identical to *w* until time *t-1*, just before the first sentient being would have

[39] Again, for the sake of argument, assuming that 'all ravens are black' is a law.

gazed upon a purple object (leading to the first experience of that shade of purple), *w** is destroyed by the Sun? Would the counterfactual 'if the Sun had not destroyed the Earth, the sentient being would have experienced purple' hold?

Tooley asserts that the counterfactual should hold in *w**, but according to the regularity theorist, 'the counterfactual cannot be true unless the appropriate psychophysical law obtains'. The fact that the counterfactual should hold, though, is a mere intuition on Tooley's part. There is no reason why the Humean must have this same intuition. For the Humean *w** would have different laws, so why *should* the counterfactual be true. His argument is grounded by anti-Humean intuitions, so perhaps Tooley's argument shouldn't overly concern the Humean.

3.6.2 Counterfactuals *with* Instances

As we have seen, a regularity theory law is constituted by its instances; there is nothing more to the law 'all ravens are black' than the total population of ravens and their colours. But all the actual ravens being black does not *metaphysically* necessitate that, were there an extra raven, it too would be black.

The necessitarian would argue that, given this is the case, there is no reason why were there to be an additional raven, it would be black, as for the Humean the laws supervene on their instances and not vice versa. In other words, if there was an extra raven there's nothing stopping it from being red, despite it being a law (as it stands) that all ravens are black. Again though, this argument from the necessitarian simply assumes that the Humean has no viable means of supporting such counterfactuals, relying on his intuition that laws govern instances. Hume, however, gives us no indication of how we should deal with this problem. We shall see when we look into Lewis's more sophisticated Humeanism, though, that a viable means of judging counterfactuals can be found.

3.7 Accidental versus Non-Accidental Regularities

The 'all men called William Shakespeare are playwrights' example came out as a law because it was a single-case uniformity. This is also an example of an accidental regularity, but there are many accidental regularities that are not single-case uniformities. Reichenbach famously gives the example that all lumps of gold are smaller than a cubic mile (there are many lumps of gold!), contrasting this with the fact that all lumps of plutonium are smaller than a cubic mile. The former is accidental, the latter is non-accidental, as a lump of plutonium that large would greatly exceed its critical mass. This problem of accidental regularities poses a serious problem for the naive regularity theorist. A law must be more than just a regularity, as without an extra condition it's impossible to pick out the genuine laws from the accidental regularities. The naïve regularity theory of laws of nature, it turns out, is subject to some insurmountable objections. This does not, however, rule out all Humean metaphysics.

3.8 Conclusions

In this chapter I have outlined the most basic version of the regularity theory of laws and the most common objections to the thesis. Although, as I demonstrated, the problem of induction is no more problematic for the Humean than it is for Armstrong[40], the problems posed by single-case uniformities, non-existent subjects, counterfactuals and accidental regularities seem to rule out the naive regularity theory of laws of nature as a plausible Humean account of laws. As in Hume's metaphysics of causation an instance of causation is just an instance of law, the problems with the Humean naive regularity theory of laws are also problems for his account of causation. However, there have been attempts to 'sophisticate' the regularity theory to overcome these problems. The sophisticated version is the topic of the next chapter.

[40] I shall later argue that it is no more problematic for the Humean as it is for the other necessitarian theories, either – but this is work for another chapter.

Chapter 4: Sophisticated Regularity Theory

4.1.1 David Lewis on Causation

Lewis (1973) provides a counterfactual analysis of causation that avoids discrimination as to which events are generally considered to be causes (that is, the *inus* conditions we would normally take to be the cause, rather than other causally relevant factors), and concentrates on an unselective 'broad and non-discriminatory concept of causation' (Lewis 1973: 162). Although Lewis assesses his counterfactuals through 'comparative over-all similarity between possible worlds' (*ibid*), by outlining Lewis's counterfactual account, the strong link between cause and law that I advocate should again become apparent.

In chapter 2 I concluded that when one is looking for the cause, X, of an event, Y, one should ignore the set of minimally sufficient conditions that were possible, but not actualised, and we should consider only whether X 'in the circumstances' was non-redundant and sufficient for Y. Lewis's account provides us with a means for evaluating whether X was necessary and/or sufficient for Y, and hence a means of judging whether X was the cause of Y. Crudely, if X occurred and Y occurred in our world, to determine whether X was the cause of Y, his method requires one to consider the closest possible world to ours where X does not occur, and consider whether Y occurs in that world. If Y does not occur in the closest possible non-X world, then X is determined to be the cause of Y.

Whether or not Y occurs in this closest of possible non-X world depends on how that world would evolve in accordance with the laws of nature of *our* world. On the face of it, this analysis may look to be unsatisfactory for those who wish to deny the reality of other possible worlds, but as Van Fraassen states, 'one may tacitly read 'world' as a 'model of our language'' (Van Fraassen 1989: 45) and get the same

results; so Lewis's analysis works equally well if we consider possible worlds to be simply complete descriptions of the way the world *could* be.

Lewis thus presents his counterfactuals and counterfactual dependencies as follows:

1. '*A* □→ *C* (if A were true then C would be true) is true at a world (*w*) iff (1) there are no *A*-worlds (in which case it's vacuously true), or (2) some *A* world where *C* holds is closer (to *w*) than is any *A*-world where *C* does not hold' (Lewis 1973: 164).
2. Although there need not be one or more closest possible world, if there were *A* □→ C would be non-vacuously true iff '*C* holds in all the closest *A*-worlds' (*ibid*).
3. If *A* is true in the actual world, our world is the closest A-world. 'So A □→ C is true iff C is. Here A □→ C implies the material condition A⊃C' (*ibid*)..

It is also worth noting that there can be counterfactual dependencies upon large families of alternatives, as is often the case in measurements. For example: Rs (the family if possible propositions R1, R2, R3...) depend on Ps (The family of possible propositions P1, P2, P3...) if the Rs are the possible readings of a barometer and the Ps are the possible corresponding air pressures (see Lewis 1973: p165).

4.1.2 Causal Dependence

For Lewis, causal dependence is to be judged through considerations of counterfactual dependence, but this has been characterised in terms of propositions, whereas *causal* dependence holds between events. This is unproblematic, however, as 'to any possible event *e* there corresponds the proposition *O*(*e*) that holds at all and only those worlds where *e* occurs. Counterfactual dependence among events is simply counterfactual dependence among the corresponding propositions' (Lewis 1973: 166).

He concludes that where neither events *e* nor *c* occur, *e* depends causally on *c* iff *e* would have occurred if *c* had occurred; and if both *c* and *e* occur then *e* depends causally on *c* iff, if *c* had not occurred then *e* would not have occurred.

4.1.3 The Transitivity of Causation

For Lewis, 'causal dependence among actual events implies causation… but causation must always be transitive whereas causal dependence may not be; so there can be causation without causal dependence.' (Lewis 1973: 167). This distinction is important as it is supposed to deal with counter-examples to the counterfactual account where, in the closest possible non-*c* world, *e* would still have occurred despite it clearly being the case in our world that *c* was the cause of *e*. If Bob took suicide pill X from a pot of pills and died, it would be absurd to claim that X hadn't killed him because if he hadn't have taken pill X he would have taken pill Y from the same pot, and that would have had the same effect. X clearly killed Bob despite the counterfactual 'if Bob had not taken X Bob would not have died' being false. By introducing transitivity into causation, the causal chain can be identified between Bob taking the specific pill X and Bob's death, where each event in the causal chain is counterfactually dependent on the event immediately prior to it in the chain. X can then be identified as the cause of Bob's death[41].

4.2.1 Laws in Lewis's Conditional Analyses

Lewis has presented a conditional analysis of causation close to the conceptual analysis provided by Mackie, in which he, too, argues that causal dependence should be analysed in terms of counterfactual dependence. Lewis's approach looks plausible, but given that his metaphysical assessment of counterfactuals is carried out by considering which possible worlds are closest, how does Lewis's (counterfactual) conditional account of singular causation rely on laws of nature? The answer to this lies both in the assessment of what considerations come into play when identifying the closest possible worlds, and the criteria Lewis

[41] There has been further discussion of problems with Lewis's account, focussing in particular on problems arising through cases of pre-emption. I shan't enter into this minefield, but for a good account see Peter Menzies' artical in *The Stanford Encyclopedia of Philosophy* (http://plato.stanford.edu/entries/causation-counterfactuals/)

imposes when deciding what events occur when the possible worlds are left to 'run on'.

So how do we determine which of the possible non-X worlds is the closest? My hypothesis would certainly be supported if the laws of the closest non-X world, *w**, must be the same laws as those of the actual world - but Lewis must deny this, so he claims that similarity of laws is 'weighty, but not sacred'. The similarities of particular fact between worlds must also be taken into consideration, as an exact match over a large spatiotemporal region is also an indication of closeness. 'It may be worth a small miracle to prolong or expand a region of perfect match' (Lewis 1973: 164). Given Lewis's account of laws, a small miracle (judged by the laws of our world) could lead to very different laws in this close possible world.

Although having the same or very similar laws in a world contributes positively to the closeness of that world to the actual world, it is not necessarily the case that the closest possible non-X world will be a world with the same laws as the actual world. As it turns out, in the closest possible non-X world, *w**, the laws of our world will almost certainly need to be broken in *w** for X not to occur, but – and herein lies the rub - '*the (counterfactual) analysis requires that (w) evolve thereafter in accordance with the actual laws*'[42] (Lewis 1986: 43). Lewis's account of singular causation in terms of counterfactual conditionals is, I think, very appealing, and laws remains absolutely central in determining which events are causal and which are not. If Lewis's account of laws of nature turns out to be unacceptable, then so will his account of causation.

Lewis holds what is sometimes known as The Sophisticated Regularity Theory (SRT), or the 'web-of-laws account' (see Psillos: 2002), which retains the non-necessitarian nature of the naïve position, whilst dealing with some of the major objections against naive regularity theory we came across in the previous chapter. The greatest benefit of

[42] My emphasis

SRT over the naive version is arguably its ability to distinguish accidental from non-accidental laws, but another benefit is that the practice of discovering the laws in the SRT matches nicely with scientific practice. Before the SRT can be spelt out in more detail, though, I must outline the principle upon which Lewis's entire metaphysics is based – Humean Supervenience.

4.2.2 Humean Supervenience

Lewis once claimed that his life's work could probably be summed up as a defence of Humean Supervenience. This is the claim that everything contingently true of a world like ours (a temporal-parts-world) is true in virtue of the patterns of fundamental property instantiation (and the fundamental relations that hold between them) in the mosaic discussed in the previous chapter – as Bigelow puts it, 'truth supervenes on being' (Bigelow 1988: 132). The doctrine of Humean Supervenience is 'that all else supervenes on the spatiotemporal arrangement of local qualities throughout all of history, past present and future (Lewis 1994: 474).' For Lewis, then, all contingent truths are true in virtue of these patterns of fundamental property instantiations, and the fundamental relations between these instantiations, and in virtue of these alone. It is thus a distinctly Humean thesis; that is, one devoid of metaphysical necessity.

4.2.3 Lewisian Properties

Lewis's properties are quidditistic; that is, their identities are not fixed by their causal/nomological roles, but by the quiddities they primitively possess. Lewisian properties are therefore categorical.

For Lewis, properties are sets of individuals[43]. The property of being red is the set of all the red things in this and all other possible worlds, and to instantiate the property of redness is to be a member of this

[43] In Work on a New Theory of Universals (Lewis 1983b) Lewis defines properties as *classes* of individuals, but he makes a point of rejecting this in favour of *sets* of individuals in *On the Plurality of Worlds.* See Lewis *On the Plurality of Worlds* Blackwell Publishing 2001 (first published 1986) pp48-53

set. As a modal realist Lewis is able to use this ontology of properties to account for contingent truths. It is a contingent truth that all ravens are black because there are non-black members of the set of ravens (as the set of ravens includes not only the ravens of the actual world, but also the ravens that exist in other possible worlds (many of which are not black)).

As there are infinitely many possible worlds, there are infinitely many sets of things. Some of these sets include only three things in our world, some include trillions across many possible worlds. Some members of properties (sets of objects) bear very little or no resemblance to one another (take the set of me, a brick, and a quark, for example). Properties are thus abundant (just think about how many sets of things there could be when you take the objects in this and all other possible worlds. It is inconceivable!). But most of these abundant properties are not the kind of properties we are accustomed to talking about – properties like colours, charge, mass etc... Lewis calls the properties we usually consider to be the natural properties, 'sparse' properties. 'Sharing [sparse properties]... makes for qualitative similarity, they carve at the joints, they are intrinsic, they are highly specific, the sets of instances are *ipso facto* not entirely miscellaneous, there are only just enough of them to characterise things completely and without redundancy' (Lewis 2001: 60). The property of charge -1, then, is still just the set of all the things across possible worlds with charge -1, but it has the special status of being sparse.

Sparse properties (individual sparse properties as specific sets of individuals) are members of the set of all properties (the set of all possible sets of individuals), but they are special in the sense of being the kind of properties we use in science and ordinary language. Lewis writes that 'we need no other entities [over and above the abundant properties], just an inegalitarian distinction between the ones we've already got. When a property belongs to this small minority, I call it a *natural property*' (Lewis 2001: 60). Sparse properties, then, are just properties that, in a sense, glow in the dark – there *is* some special significance to them, but they sets of particulars just like the abundant properties. There is a further distinction

to be made, however. The property of being a raven is a sparse property (on the face of it any property we have a name for is likely to be sparse). However it is not as natural as the property of having charge -1. Lewis thus allows for degrees of naturalness. Some properties - the ones that all and only their members would instantiate a single universal, were a theory of universals tenable - are *perfectly natural*. (See Lewis 1983: 344-346) For Lewis, the *real* universals (were universals a tenable theory of properties) would be those properties of fundamental physics.

As we saw, Lewis names some of the qualities of natural properties, but one may question what *really* distinguishes these from the very unnatural properties; that is, is there anything in their *nature* that makes them more or less natural? The answer is a simple no. Lewis's is a nominalist account: according to Lewis, whether a property is unnatural, natural or perfectly natural is primitive matter, where either the predicate 'natural' is itself primitive, or if this is unacceptable, perhaps what distinguishes them is 'primitive objective resemblance among things'(*ibid*); either way, there is no special metaphysical distinction between natural and unnatural properties.

4.3.1 Lewis's Best Systems Analysis of Laws

In the preceding chapter I outlined some of the major difficulties facing the naïve regularity theorist. The major unresolved issues arose from the problems found when trying to support counterfactuals, and when trying to distinguish laws from accidental regularities. Lewis makes important changes to the regularity theory of laws, whilst retaining its non-necessitarian ideal, in order to solve these destructive problems. Along with Mill and Ramsey, Lewis develops a best systems analysis of laws. This system is also sometimes known as the web-of-laws (Psillos 2002: 137-159), or the Mill-Ramsey-Lewis account. Mill writes:

> According to one mode of expression, the question, What are the laws of nature? May be stated thus: What are the fewest and simplest assumptions, which, being granted, the whole existing order of nature would result? Another mode of stating it would be thus: What are the fewest general propositions from which all

> the uniformities which exist in the universe might be deductively inferred? (Mill 1911: 207)

Lewis construes it slightly differently. Under Lewis's account, the ideal set of laws would be the set of regularities holding between *perfectly natural properties and relations* that (a) provides all the information about the world - entailing all the world's truths; and (b) is the simplest possible set of regularities.

Given that the properties in these regularities have to be perfectly natural, satisfying both (a) and (b) is, in most cases, impossible[44]. *All* the information about the world could be gathered only if all the world's truths were laws, and this surely cannot be right. Similarly, the simplest regularity would provide us with only a tiny fragment of information about the world. Given the impossibility of satisfying both (a) and (b) in most worlds, the best systems analysis cannot ask for this, and so Lewis concludes that a 'contingent generalization is a law if and only if it appears as a theorem (or axiom) in each of the true deductive systems that achieves a best combination of simplicity and strength' (Lewis 1973: 73). The laws turn out to be the regularities holding between perfectly natural properties and relations between them derivable from this ideally simple and strong (and simple and as informative as possible) description of the world. The parameters of Lewis's conception of laws are therefore far better defined than that of the naive regularity theorist.

Unlike the naive regularity theory of laws, the best systems analysis prevents any old regularity counting as a law. Accidental regularities like 'all people called William Shakespeare are great playwrights' will not come out as laws of nature - firstly because it would not form part of or be derivable from the best system, and secondly because 'being called William Shakespeare' is not a perfectly natural property. Regularities involving 'gruelike' properties are also ruled out as they are primitively unnatural.

[44] Perhaps both these criteria could be satisfied in some very uninteresting, simple worlds.

Not only do the fundamental laws or axioms of the best system count as laws, but so do any propositions that logically follow from these axioms. On the face of it this is great, as many of the statements we accept as law-statements do not express the basic laws, but something less fundamental. The statement 'all electrons have charge -1' does not express one of the fundamental laws in our world[45], but it is widely accepted as a law nonetheless. To deny this law-status would at the very least be counter-intuitive. Unfortunately we probably end up with far more laws than we'd like, but this is a bullet the sophisticated regularity theorist must just bite.

4.4 Does the Best Systems Analysis Solve the Problems of the Naive Regularity Theory?

In this section I turn back to the remaining knock down objections to Hume's naïve regularity theory (namely, the problems posed by non-existent subjects, functional laws, and counterfactuals with no instances) to see whether Lewis's account fares any better.

4.4.1 Non-Existent Subjects

One of the problems with taking laws to be universal generalisations without further restrictions was that we could end up with an infinite number of incompatible laws. Where a law takes the form $\forall x(Fx \rightarrow Gx)$, if there are no Fs in a world then the proposition is true for any G. I gave the example that it could be a law that 'all unicorns are 2 metres tall' and a law that 'all unicorns are 3 metres tall' in the same world. But this cannot be true once we impose the sophisticated regularity theorist's objections. The best system must be a coherent set, and so it's logically impossible for regularities derivable from that set to contradict one another.

[45] We're unsure what these fundamental laws are, of course, but we're looking to unify quantum mechanics and general relativity to form one uniform theory of quantum gravity. This would be a regularity from which all other laws could be derived. Even if there is no such regularity, laws like all electrons have charge -1 are not going to be fundamental in our world

One interesting point to note, however, is that regularities with no instances *can* appear as laws in the best system. It is not logically impossible (in fact it is probably quite common) for laws without instances to be derivable from the best system.

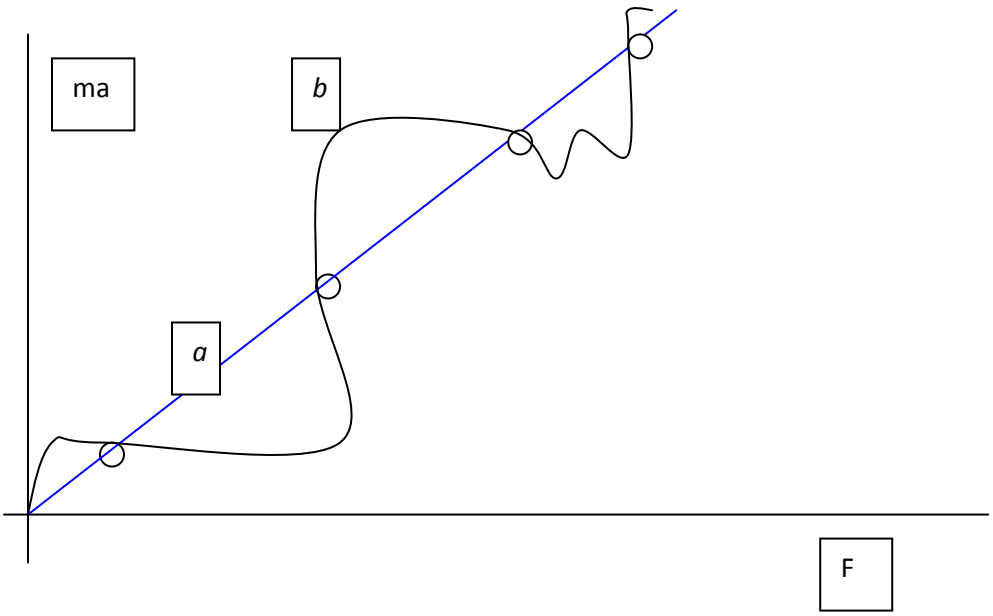
4.4.2 Counterfactuals with no instances

In section 3.6 we looked at Michael Tooley's objection to the naive regularity theory, in which he asked us to consider two worlds (*w* and *w**) identical up until time *t-1*. In *w* at *t* a psychophysical law connecting neurophysiological states to phenomenological states determines that an individual gazes upon a purple object and experiences the phenomenological state 'purple'. No purple object had ever been observed before. But in *w** no individual ever gazes upon a purple object, because the Sun destroys the Earth at *t-1*. At *w** the psychophysical law that obtains at *w* is not derivable from the best system, and so it seems the counterfactual 'if the Sun had not destroyed the Earth, the sentient being would have experienced purple' does not hold.

This looks problematic for Lewis. If the psychophysical 'law' is not derivable from the best system at *w*, then it is not a law at *w*, and Tooley would be right. However, as I suggested in 3.6, I think Lewis can just bite the bullet here, and claim either that (a) nature would have to be very unkind for this law not to be derivable from the best system given that the local particular matters of fact match perfectly at *w* and *w** until *t-1*, and (b) if it turns out that this micro-physical law is not a law, then so be it. A Humean need not find this outcome overly counterintuitive.

4.4.3 Functional Laws

As we saw, although this is not explicit in Hume's work, the naive regularity theorist would have to deal with functional laws by invoking something like primitive simplicity (as given just the rules of the naive version, we have no reason to assume the right function to be inferred below is the straight line, or the squiggly one).



Looking at the graph we can see that all the nodes are in a position that would satisfy the equation F=ma (denoted by *a*), but they would also satisfy whatever equation would be required to draw the function denoted *b*. Whereas the naïve regularity theorist has to make *ad hoc* additions to his thesis to choose *a* over *b*, it is very likely that for Lewis *a* would always come out as the function we must choose, as (unless nature is unkind) only *a* will ever be derivable from the best system. Using *a* and the criteria for closeness of possible worlds, the sophisticated regularity theorist can now support the counterfactual 'if there had been a force of 5 Newtons acting on an object of mass 1,000,000kg, then that object would have accelerated at 5 $x10^{-6}m/s^2$'.

I have hopefully now shown that the objections raised against the naive version of the regularity theory do not pose so much of a threat against SRT. However there are further objections more specifically aimed at SRT that need to be addressed.

4.5 New Problems for the Humean

Although the sophisticated Humean overcomes the problems faced by the naïve regularity theory, the SRT faces a number of additional objections:

4.5.1 The laws are 'psychologistic'

As the best systems analysis is grounded on laws being derivable from the systems of regularities 'achieving the best combination of simplicity and strength', it would be best if we had some objective, language-independent way of measuring and finding the right balance between them. If no such standard exists, then different standards will sometimes come up with different fundamental laws. What the true law-statements are would become dependent on *our* interpretations of simplicity and strength and the balance between them, but one of the main criterion we generally want to uphold is that laws should be mind-independent. According to Daniel Nolan, 'Lewis does not tell us what [the] objective standards are, but he claims there are such standards to be discovered' (Nolan 2005: 85). If Nolan is correct then the 'laws turn out to be psychologistic' objection doesn't work, but in Lewis's own words suggest he *does* accept the possibility that there might be two equally reasonable standards of the balance between simplicity and strength. However, he concludes that 'if nature is kind to us, the problem needn't arise. [He supposes] our standards of simplicity strength and balance are only partly a matter of psychology... if nature is kind, the best system will be *robustly* best... I'd blame the trouble on unkind nature, not on the analysis; and I suggest we not cross these bridges unless we come to them' (Lewis 1994: 479). Of course in some worlds, these bridges may need to be crossed!

4.5.2 Systems that are 'Equal Best'

There is a similar objection which holds even if there *is* just one objective standard of simplicity and strength and balance. It is possible, although unlikely, that two distinct systems of regularities might come out as equal best on this objective scale. In which case choosing between one set of laws and another would again be a purely psychological matter. We surely cannot conclude that both systems provide us with the laws of that world, because one set of laws could determine some regularities to be laws that

the other rejects as laws[46]. Lewis's response is simply that we would have to concede that the laws coming out of the 'barely-best systems would not very well deserve the name of laws'(*ibid*), but somehow this feels unsatisfactory when we're looking for a wholly objective metaphysical analysis of laws. Nevertheless, I think that Lewis's thought that science is extremely unlikely to give us contradictory best systems even if we use different standards is probably right. Given the confidence with which we can make this assumption, trying to work out what these standards are at this time is unnecessary. As far as I'm concerned, although it is hidden from us, we may as well take it that there is a primitive objective standard that can be applied, and furthermore that in most worlds, this standard will provide us with only one set of laws.

4.5.3 The laws do not determine the instances

Once again we come across the objection that laws are supposed to *metaphysically determine* the regularities and their instances, not the other way around. If, as the objector would claim, laws must govern the way things in the world behave[47], then the best systems analysis does not give us laws, as they clearly do not satisfy this condition. In contrast, Armstrong's natural necessitation relation between universals plays precisely this role: 'This F is a G' would be metaphysically determined by the pre-existing necessitation relation holding between Fs and Gs (the law). If we want Lewis's laws to determine their instances, then the instances of Fs being Gs would need to be metaphysically determined by the proposition 'all Fs are Gs', and of course a proposition cannot determine a state of affairs.

However, it seems to me that the sophisticated Humean need not deny that instances of laws are determined. What they would deny is that the term 'determined' requires there to be *metaphysical* necessity - F is

[46] Lewis originally thought that the laws would be the law-regularities common to the best systems of all the reasonable standards for simplicity strength and balance (Lewis 1973: 73), but he later rejects this view (Lewis 1994: 479).

[47] 'Governing' here can be taken as 'metaphysically necessitating', although they need not be taken as equivalent (as we shall see when we look at the dispositional analyses).

physically determined to be a G if 'all Fs are G's' is a law of nature, and the Humean has an account of laws. It is true that 'this electron has charge -1' is partially determined by electrons millions of light years away and thousands of years in the future having charge -1, as the law itself depends upon these facts, but, says the Humean, so what? As Beebee writes, 'laws *are*, in part, facts about the future' (Beebee 2000: 578). When the anti-Humean says that laws determine the events, they are ultimately talking about a different kind of determination: one that requires metaphysical necessity. When they object that the best systems analysis of laws has no metaphysical necessity, no Humean will be troubled, as Humean Supervenience tells us that all truths in our world supervene on the spatiotemporal arrangement of local qualities. All the objection amounts to is an expression of anti-Humean intuition (and presumably Lewis does not have these intuitions!).

4.5.4 We Can Neither Discover nor Use the Laws

It is also argued that given Lewis's conception of laws, we cannot discover what the laws are, nor use them in predictions. I hope I dealt with this objection in chapter 3 when offering a defence of Humeanism against the problem of induction. Ultimately, Armstrong and other necessitarians have access only to the same information (regularity observations) as the Humean, so they have to infer their laws inductively, too. If they wish to invoke inference to the best explanation to do this then that's fine, but to use it as an objection to Lewis they have to show that their explanation of the observations is a better explanation than the Humean's, and given the discussion in section 3.3, we have good reason to think otherwise.

4.5.5 Counter Examples

Tooley, Carroll and Menzies (cited by Beebee 2000: 584) try to refute the best systems analysis by providing a counter example. They attempt to provide us with two very simple worlds that precisely match in local particular matters of fact, but obviously differ in their laws. Each world has (only) one particle and one field (call the particle in world *w* X1, the

field in *w* Y1, the particle in world *w** X2, and the field in *w** Y2). In *w* we are asked to accept that it is a law that all X-particles, when they enter a Y-field, have spin-up, but in *w** , we are told, it is a law that once an X-particles enters a Y-field, it has spin-down - but neither X1 nor X2 ever enter into the Y-fields of their respective worlds. If this situation is possible, then the best systems analysis must be wrong, as *w* and *w** *must* have the same laws as their 'mosaics' match precisely.

But this counter example does not work. If we take the complete description of *w* and *w** to include its laws, then '*by its own lights*: it isn't as if the Ramsey-Lewis view entails that *w* and *w** are possible' (Beebee 2000: 584). To re-state my defintion in 4.2.1, the laws are regularities holding between perfectly natural properties and the relations between them derivable from the ideally simple and strong (and simple and as informative as possible) description of the world. The vacuous laws proposed by Tooley, Carroll and Menzies do not qualify as laws in this world under Lewis's account, as they are not derivable from the ideally simple and strong description of the worlds.

4.6 Categorical Properties and Quidditism

In previous chapters I have said little about categorical properties except that these are non-powerful, non-potent, non-dispositional properties. They provide no driving force for particulars to move, they provide no 'necessary connections', 'oomph', 'biff' or 'metaphysical glue' (or any other of the seemingly endless list of what I consider simply to be virtual synonyms in the necesssitarian vocabulary). There is more to say about categorical properties, however, than what they are *not*, so in this section I shall give a positive account of what it is to be a categorical property.

A categorical property is a property whose nature is quiddistic[48]; so the identity of that property is a brute fact, not fixed by its causal or nomological role, but by the quiddity it possesses. 'Paradigmatically, a

[48] In section 7.3.3 I will explore the possibility that there are some categorical properties with non-quiddistic natures.

categorical property is thought to be a property whose identity is fixed by a quiddity' (Barker 2009: 1).

If we accept quidditism, 'the acceptance of primitive identity between fundamental qualities across possible worlds' (Black 2000: 92), it is a primitive fact that property F plays dispositional role[49] D in the actual world, and in many other possible worlds property F will play entirely different dispositional roles, or perhaps even more than one dispositional role. In the closest possible worlds, F will play the same role as it does in the actual world, as we look for match of both laws and particular matters of fact, but there will be more possible worlds in which F does not play the D-role than worlds in which it does.[50]

4.6.1 Quidditism in Sophisticated Humeanism

As we have seen, for Lewis a property is a set of spatiotemporally located instances, where its instances are members of the property. Given his Humean supervenience, Lewis's laws supervene upon local particular matters of fact, and the dispositional role of each of these local particular matters of fact is not fixed outside of the metaphysically contingent laws that hold in virtue of the Humean mosaic.

Where a property is a set of particulars, there are some properties whereby all members of the set in the actual world play the 'blackness' role, but in other possible worlds they play the 'charge -1' role. Lewis's properties are thus quidditistic. The property F is the set of all the actual and possible objects in the F-set, whose dispositional role in any world is a brute fact determined by the Humean mosaic. F could therefore play the blackness-role in the actual world, *w*, and the redness role in *w**. In fact, the property F need not even be a colour in other possible worlds - it could play the dispositional role of 'making a noise when scratched', or any

[49] 'Dispositional' here takes the conceptual meaning; it refers to the kind of causal role the property plays but implies no metaphysical glue.

[50] Note there will be an infinite number of worlds in which F plays the D-role and an infinite number of worlds in which it doesn't – so perhaps this 'many more' is the same kind of claim as there being more natural numbers than odd numbers. This is unimportant for the purposes of this chapter, however.

other role for that matter. Similarly, the property G that plays the dispositional role of gravitational attraction in our world could, in *w**, be the property playing the dispositional role F plays in our world.

In light of these implications Mumford writes that '[Having a quiddity] allows that F and G could swap their entire causal roles and yet still be the same properties they were' (Mumford 2004: 104), we also see in Black: 'Lewis explicitly allows... a world isomorphic with ours, but where one of the quark colours has traded place with one of the flavours' (Black 2000: 92), and again in Bird : '(QA1) For all fundamental universals F and powers X there is a world where F lacks X' (Bird 2007: 71). As Bird points out there is no obvious logical contradiction in holding the view that properties can swap powers, so the objection must be simply that this is extremely counter-intuitive. But counter-intuitiveness is not necessarily a proof that the thesis is wrong.

Bird admits that counter-intuitiveness alone does not prove categoricalism about properties false, even if it does motivate the search for another option. In light of this he provides a further objection that he believes shows the thesis to be undeniably false. It seems to be a further consequence of properties having quidditistic natures that in the same world two distinct properties could play the same dispositional role. This looks to follow from categoricalism, as, if a property's identity is a brute fact (and in no way fixed by its dispositional role), there is nothing to stop two properties, F and G, playing the same role in the actual world (two distinct sets can share members).

Bird writes that according to quidditism, 'One and the same world *w* is such that (i) at *w*, universal F has powers {C1, C2,...}; (ii) at *w*, universal G has powers {C1, C2...}; (iii) F≠G' (Bird 2007: 76). In this world it would follow that it were a law that 'All Fs are Hs', it would also be a law that 'All Gs are Hs'. This is problematic, Bird suggests, because where it is a law that all ravens are black, where both Fs and Gs are universals denoting ravens, and H is the universal playing the blackness-

role[51], when we see a black raven 'there are two possible explanations for this': all Fs are Hs, and all Gs are Hs. We could never know which – it could even be both! Bird argues that given this possibility of multiple realisability, our term 'raven' can never refer (See Bird 2007: 76-77).

4.6.2 Against Against Quidditism – Another Humean Position?

I would like at this point to present a tentative response on behalf of the nominalist quidditist. The quidditist may, on the face of it, be able to claim that Bird's (QA1) is false of what he calls fundamental universals. We saw that Lewis makes the distinction between sparse and abundant properties – unnatural, natural and perfectly natural properties. One nominalist account of the distinction between natural and unnatural properties is to say that the perfectly natural properties are the sets of individuals wherein their members across all possible worlds are all and only those particulars that are primitively qualitatively similar in the relevant respect[52]; another is to take the sets of natural properties to be primitively natural (see Lewis 1983b: 344). 'Natural properties would be the ones whose sharing makes for resemblance, and the ones relevant to causal powers' (*ibid*).

Properties are sets of their instances across all possible worlds, and the natural properties are those usually associated causal powers – these sets, for Lewis, are glow in the dark sets, although we are never really enlightened any further. Here are three possible ways in which we might associate a property with a causal power:

(i) The perfectly natural properties are sets whose instances in our world each play a role associated with a single causal power in our world, where all the property's instances are located in our world.

(ii) The perfectly natural properties are the sets of instances associated with a single causal power in our world, *and* instances associated with single causal powers in other possible worlds, but perfectly

[51] My example.
[52] Lewis acknowledges the technical difficulties with this, but claims they can be solved (albeit at a 'daunting price of complexity and artificiality of our primitive)

natural property F in our world may be associated with a different causal power in our world than it is in other worlds. (So F might play the charge role in our world but the mass role in another. Both are relevant to causal powers.)

(iii) The perfectly natural properties are the sets of instances associated with causal powers in our world, and *the same* causal powers in other worlds.

Now if either (i) or (ii) were the correct interpretation, then the set of objects across all possible worlds that all play the charge role (which is of course a property) would not be the natural property corresponding to the charge role in our world. This seems odd as it would indeed allow for the permutation of a property's causal role across worlds, but nonetheless possible given that the identity of properties is not fixed by their causal roles.

Lewis states the relation between instances of a property could be one of primitive resemblance - but what is meant by primitive resemblance? Given what we are trying trying to achieve by invoking the naturalness of properties (which is essentially to give us the properties we can use in laws), why not think any one member of a property resembles any another in the way we would normally use the term; that is, it is a primitive fact that the objects constituting a natural property look similar/behave in similar ways. Now, the resemblance of the members of these properties stretches across possible worlds, so why restrict our interpretation of resemblance to the actual world? Why shouldn't we allow the Humean to say that the instances of a perfectly natural property in our world also resemble (in our sense of appearance and causal role) its instances in other possible worlds? Lewis-style glow in the dark sets turn out to be the sets whose members are all and only the objects, actual and possible, whose members primitively resemble one another. This is not to say that it is *in virtue* of their primitive resemblance that they are the natural properties – it just happens to be the case that the natural properties

are these sets. The identity of the natural properties is fixed primitively; that is, by a quiddity.

Under this quidditistic view, it is not possible for P to be the perfectly natural property corresponding to the property of charge -1 if it plays the repelling protons-role in the actual world (the charge -1-role), but the attracting protons-role in another, for this would violate the resemblance of the members of the property in the charge-respect. Lewis tells us that 'sharing of the [sparse (or natural) properties] makes for qualitative similarity... and there are only just enough of them to characterise things completely and without redundancy' (Lewis 1986b: 60). If we allow the identity of natural properties to be fixed primitively, but it turns out that this primitive naturalness coincides with brute *qualitative similarity*, the quidditism arguments presented against Lewis would not hold much weight (at least not at the fundamental level). Bird argues that we have two explanations for all ravens (Fs and Gs) are black (H) – all Fs are Hs, and all Gs are Hs – but given (a) that the explanation must be a law, and (b) laws only include natural properties, only one of those explanations can be the right one, as only one of them is a law (the one only including natural properties). Our term 'raven' refers – it refers to the natural property: 'raven' – that is, the only set of all the actual and possible ravens.

In short, given the nominalist exposition of the nature of a natural property, the nominalist might argue that it does not make sense to claim that a natural property-term does not refer (as more than one property play each 'natural-property-role'), because it is a *de re* necessary truth that only one of them can be a natural property; that is, the set of all the actual and possible particulars that resemble one another in that particular natural property-respect. Before applying the objections presented in the previous section, then, the proponent of the quidditism objection must first show that Lewis cannot draw the distinction he does between natural and unnatural properties as I've presented it, and it's not obvious that the proponent of this objection can.

A number of issues will be raised here, though: firstly, to escape this quidditism objection my Humean is linking natural properties to causal roles (this is ultimately what happens when we start talking about brute similarities). Why should this be the case? The Humean can answer only that it is a primitive fact. Again, this is not a knock-down objection, but nonetheless we are introducing more and more scary primitives into this Humean metaphysic. At least, though, we have replaced the spooky 'glow in the dark sets' with what I consider to be the more attractive 'primitive similarity'.

One might claim that linking a property to its causal role undermines the entire Humean project. In all possible worlds, if an object has charge -1, then it repels objects of charge +1. The whole point of Humeanism seemed to be the denial of this fact. However, it seems to me that at its core, Humeanism is about the denial of necessary connections in the objects – a denial of metaphysical glue, of 'oomph' and 'biff', and so on. This characterisation of Humeanism still does not have any of these components in its ontology. It still has just particulars, and sets of particulars. There is no glue.

The necessitarian might also respond that without invoking universals or tropes, this position is still quidditistic (which the Humean would embrace, incidentally), and by the very nature of quidditistic ontologies, causal-role permutation for properties is possible. They will argue that you cannot take natural properties to be those that do not permute their causal roles across possible worlds, as this violates quidditism about properties. But this, I think, is to confuse what quidditism is all about. Quidditism is a thesis about what fixes the identity of a property, not about causal-role permutation, whereby the nature of a property's identity is fixed primitively, and is entirely independent of modal facts. This is true of all ontologies in which properties are sets of particulars. One cannot look at the abstract entity that is a set, and from its nature tell what its powers are. The identity of the property is thus still a

primitive fact, and it is the fact that the members of the glow in the dark sets are primitively qualitatively similar.

4.6.3 *De Dicto* Responses

Bird presented his arguments as an ultimate refutation of quidditism - but Harold Noonan has recently published a paper showing Bird's argument to fail due to a conflation of *de re* and *de dicto* necessary truths (Noonan 2010). Bird claims:

> (QB2) One and the same world *w* is such that (i) at *w* a property F has powers C1, C2, ...; (ii) at *w* property G has powers C1, C2, ...; (iii) F is not a G (Bird 2007: 75)

But Noonan argues that, just as under haecceitism even though two men *x* and *y* might swap all their accidental properties (whilst keeping the same essential properties) and retain their identity, it does not follow that *x* could have all *y*'s properties and *y* retain his own properties. This is because although *x* and *y* can retain identity whilst swapping their properties, it is a *de dicto* necessary truth that no 'two men can occupy the same place(s) at (all) the same time(s)... *Mutatis mutandis* if property's identity is independent of powers in the sense implied by (QB1), i.e., that in distinct worlds distinct properties have the same powers, it does not follow that two properties can possess all the same powers in the same world' (Noonan 2010: 78).

Bird is making a jump from the *de re* claim that where two properties (sets of particulars) play the same role in a single world, to the claim that "the property that plays the blackness role" does not refer. But as the *de re* claim is perfectly consistent with it being a *de dicto* necessary truth (and Humeans have no issues with *de dicto* necessities) that no two properties can have the same causal powers in a single world, Bird's argument is fallacious.

Chapter 5 Armstrong's Governing Conception of Laws of Nature

Armstrong's view has received much criticism. Objections have been raised that I feel successfully refute the position, so in this chapter I will first provide a synopsis of the position, and then show why it is untenable by outlining these objections. This work is important, though, as I will ultimately demonstrate that the dispositional essentialist positions collapses into a new form of Armstrongianism, and so the same arguments presented in this chapter can ultimately be used to refute dispositional essentialism.

'Armstrongianism' is motivated by a conceptual assumption: namely, that the spatiotemporal distribution of properties should be determined by *governing* laws, rather than laws supervening on their instances. Armstrong wants a metaphysical conception that solves all the problems with the Humean positions: one that explains past regularities, allows us to make inductive inferences, provides support for judging counterfactuals (all of which I claim are not problematic for the Humean, anyway), and where laws 'govern' their instances. As we have already seen, for many (all of them anti-Humeans!) there is an intuition that laws govern - in a metaphysically meaty sense - so it is worth showing why Armstrong's governing conception of laws as relations between universals fails, and thus motivates the search for an alternative anti-Humean position. In the next section, though, I shall provide a little more detail about what the Armstrongian view is, and why its proponents think it should be adopted.

5.1 The Theory

Unlike the sophisticated Humeans, Armstrongians reject nominalism about properties in favour of universals. Armstrong adopts the following views:

1. That all natural properties are universals;

2. Immanent realism about universals; that is, one and the same universal is wholly present in each instantiation of it, but it exists only in its instances. (The universal 'charge' therefore exists only where there is an instance of charge; this clearly contrasts with the view that they exist transcendentally.);

3. That there is a hierarchy of universals. Examples of first order universals would be 'red' and 'charge -1', but there are also second order relations between first order universals, and so on;

4. There are second-order natural necessitation relations, N, that hold between two (or more) first order universals, denoted N(F,G). Where N(F,G) holds in a world, it is metaphysically necessary that all Fs are Gs – N(F,G) is a law of nature.

In light of these claims, Armstrong would posit that there are no laws in worlds where there are no instances of the second-order universal, N, standing between two first order universals. This, I suppose, is intuitively appealing for those without Humean intuitions. Necessitarians may well want to claim that where there are no necessary connections between distinct existences (in the objects) in a world qualitatively similar to ours, there are no laws[53] - Armstrong's position, if tenable, provides us with the requisite metaphysical necessity to satisfy these claims.

Suppose in world *w* the states of affairs: 'N(F,G) and *x* is F' is actualised. It follows that *x* is also G. Although N, F and G are all universals whose causal/nomological roles are contingent (as N(F,G) holds only contingently in any world), it is still true that in all possible worlds in which 'N(F,G) and *x* is F' is true, '*x* is also G'. We therefore have necessary connections between distinct existences, despite the properties themselves having no fixed transworld causal role.

[53] This intuition arises from our concept of causation/law in which necessary connection plays an integral role. However as we discussed in the opening chapters, necessary connection can be a part of our concept of causation, but need not a part of our metaphysics of causation.

According to Armstrong, then, non-accidental regularities differ from accidental regularities, because laws involve a natural necessitation relation between the universals. So:

1. Armstrongianism provides a viable way of distinguishing between accidental and non-accidental regularities;

2. It provides a means of supporting counterfactuals;

3. There are explanations for uniformities. '[Unlike with the regularity theorist],the modal character of laws is explained' (Mumford 2004: 87);[54]

4. Laws govern.

If the view works, it satisfies all the desiderata for the necessitarian.

5.2 Causation in Armstrongianism

Armstrong (1983) initially claims that the link between cause and law in his view is just *de facto.* But unsurprisingly there is a far stronger connection than this. For the naive regularity theorist a cause is just an instance of a regularity (which equates to an instance of law). As we saw, this view of course failed for numerous reasons. Armstrong, however, also sees a cause-effect relation to be an instance of nomic-connection. Unlike the naive regularity theorist, the holders of the Armstrongian view do not take every regularity to be a law. Their laws are 'strong laws' (Heathecote and Armstrong 1991: 69); that is to say the only regularities that actually count as laws are those where the natural necessitation relation sits between the relevant properties.

What justifies their belief in this 'cause as instance of law' view is, they say, the repeatability of causal interactions. When we observe what

[54] I cannot agree with Mumford here. Lewis has a very clear conception of modality based on his view of laws, properties and modal realism. With SRT the laws are contingent upon the spatiotemporal location of local particular matters of fact. The 'modal character of laws' has been thoroughly explained!

we think is a causal process where event *c* looks to be the cause of event *e*, in science we try to isolate the more specific properties involved in the causal interaction. We make small adjustments to the circumstances in which *c* originally occurred and see whether *e* still occurs. To take a very simple example, if I dropped a cold cube of glass into a glass of water and the water got colder, using scientific method to determine precisely what property of the 'macro-cause' caused the water to get colder, I would repeat the experiment with small changes. I'd try dropping a cube of glass with a higher temperature into the water and re-measure the temperature of the water. I would then drop a cube of iron at the same temperature as the original cube of glass into the water and see if the water got colder, and then a spherical block of the warmer glass into the water and see if the water got colder... We could eventually conclude that the relevant property was the temperature of the substance dropped in the water.

Once these experiments have been conducted and the relevant properties isolated, we can repeat the experiment many times, and just as Kripke finds that it is *a posteriori* necessary that water is H^2O, so it is discoverable *a posteriori* that there is natural necessitation relation between the identified cause-property and the identified effect-property. As every causal interaction is an instance of one (or more) universal followed by an instance of another (or more than one) universal where the universals involved are linked by the natural necessitation relation, we can say that every causal sequence is an instantiation of a law.

It does not follow, of course, that every instance of law is an instance of causation. Instances of synchronic laws like 'all electrons have charge -1' would not be instances of causation, so extra conditions need to be assigned. Regardless of what these conditions are though, all causes are instances of laws, and so again the strong connection between cause and law is evident in the Armstrongian conception

5.3 The Supposed Advantages of Armstrong's Metaphysic

I will now take a look at the implications of the Armstrongian view as set out in section 5.2 (starting with number 4). As Helen Beebee points out, Armstrongians often claim that:

> The prevalence of the view that laws play a governing role suggests a quick refutation of Humeanism: if it is a *conceptual* truth that laws govern, then Humeanism, which accords laws no such status, must be false on conceptual grounds... According to this line of argument, Humeanism is based on a conceptual error: that of thinking that it is conceptually possible for something that does not govern to be a law' (Beebee 2000: 573).

Laws, according to the Humean, are true descriptions of the world in terms of regularities (with the various restrictions imposed by the best systems analysis), but if the concept of law *requires* laws to govern in the stronger sense; that is, one that includes metaphysical necessity, then Humean 'laws' just cannot be laws. Armstrong's natural necessitation relation, on the other hand, gives laws precisely the kind of governing role our concept of law is held to require.

It is, I think, important to note that even if this were the case, the main implication would be that the Humean ontology would be one without laws. The world could still be a Humean mosaic, where all truths supervene on local particular matters of fact - so this objection does *not* rule out the basic Humean ontology. However, there are further unwanted implications for the Humean, for without laws, the Lewisian counterfactual account of causation cannot function. Given my success theory position when it comes to the conceptual analysis of cause and law, though, if there turns out to be no laws in the world, then something has gone wrong with the analysis of the notion of laws.

The Humean must admit that their laws do not play a governing role – at least in the sense of 'metaphysically governing'. So we cannot attack the objection on these grounds. The question is, though, does the

concept of law *really* involve a governing aspect[55]. I discussed this issue in chapter 2, concluding that this particular *de dicto* argument against Humeanism was inadmissible, but to emphasise why I think it is perfectly plausible for laws not to govern their instances, we can look to Helen Beebee's (2000):

Beebee argues that this governing aspect of laws probably derives from aspects of *prescriptive* laws, unrelated to the kind of metaphysical endeavour we are currently engaging in – for example, the State governs our actions in such a way that if we break the law then we are held accountable for it. But, claims Beebee, '*a priori* reflection on the nature of *natural laws* by themselves does not yield any requirement to think of them as playing a governing role[56]'. (Beebee 2000: 582). The trouble with the 'laws must govern' objection, then, is simply that those who believe a governing role is central to the concept of law are just asserting their own 'laws must govern' intuition. Why should the Humean, who does not share these intuitions, concede that any arguments resting on the conceptual claims of the anti-Humean are at all troublesome? As we have seen, the sophisticated Humean thinks he can use his laws to support all the criteria laws need to. 'Laws govern' is not one of these conditions, and because he can provide a more ontologically parsimonious metaphysic whilst avoiding the necessitarians' appeal to some entity of which we have no positively contentful idea of, the Humean account of laws as purely descriptive entities is, according to the Humean's intuitions about laws, at least, far more appealing.

With respect to the first three 'benefits', we should remember that the sophisticated Humean satisfies them all, without postulating the N-relation. Not only that but with respect to claim 2, as I have already argued, regularities in themselves are often the best explanations for our observations.

[55] Interestingly some proponents of the dispositionalist ontology also deny this (Bird 2007: ch.9.)
[56] My emphasis

Not much headway made so far. But let us suppose (quite incorrectly!) that Armstrong's necessitation relation is somehow better than sophisticated Humeanism in so far as it provides better explanations of our observations, more of a right for us to reason inductively, and so forth. These virtues alone cannot be sufficient for us to accept the theory – it has to stand up to metaphysical scrutiny. There are important *a priori* objections that must (and, as it turns out, cannot) be resolved.

5.4 What is the N-relation, anyway – a pseudo objection?

The N-relation is a relation (a second order universal) that holds between first-order universals – the natural properties. But how are we to understand this entity? The answer is simply that we cannot. It is an unanalysible relation that exists in the world and plays the role Armstrong assigns it. To explain a phenomenon like 'all ravens are black', then, Armstrong appeals to a primitive matter of fact – that there is an unanalysable natural necessitation relation holding between the two natural properties. The regularity theorist appeals to higher-order regularities, but ultimately he too appeals to a *sui generis* fact: the relevant highest-order regularity. This is not a knock down objection to Armstrongianism by any means, but I wish to once again highlight that the proponent of the view is as much committed to primitives as the Humean.

5.5 Quidditism in the Armstrongian View

I presented the quidditism argument against Lewis in the previous chapter, and concluded that if Lewis could justify the claim that the perfectly natural properties are perfectly natural due to primitive resemblance (where primitive resemblance is relevant to resembling causal roles), then SRT is not susceptible to the *de re* quidditism objections. Furthermore, Noonan's *de dicto* argument provided significant support for the Humean. But that's not to say this response to the *de re* quidditism objections hold weight against other metaphysical theories with different conceptions of properties.

Much of what was said by the necessitarians in answer to the Lewisian view applies equally to the Armstrongian conception. As we have seen, Armstrong explains regularities by appealing to governing extrinsic laws. The particulars' properties are themselves are inert, and it is the natural necessitation relations between the universals that determine how the particulars that instantiate those universals behave. But if the universals confer no causal powers, they are categorical in a way that leads to exactly the same problems posed by dispositionalists against Humean ontologies. As the natural necessitation relations hold contingently in every world, even if N(F,G) holds in the actual world there will be many worlds in which N(F,G) does not hold, so although for Armstrong the properties are universals and not classes of individuals, there remains nothing to stop the property F playing a completely different dispositional (causal/nomological) role in another possible world than the role it plays in the actual world. But I don't think Armstrong can draw a distinction between natural and unnatural properties in terms of primitive similarity between the objects that instantiate them, as the objects instantiating these properties play no constitutive role in universals. It seems to me, then, that there would be no such thing as an unnatural universal[57], and so the causal role-permutation arguments raised against quidditism seem to work. However, one must again note that these anti-quidditism objections are nothing more than counter-intuitive implications of Armstrong's view.

Furthermore, there is no reason why Noonan's '*de dicto*' argument that Bird's proposal is fallacious, is equally useful for the Armstrongian as it is for the Humean.

5.6 The Knock Down Objection

There is, however, a knock down objection to the Armstrongian view that does not apply to the Humean. The N-relation, for Armstrong, is a second-order universal, but it is a universal nonetheless, and we know already that

[57] Note, though, that Noonan's *de dicto* response could be used against Bird by the Armstrongian, too.

universals are categorical properties for Armstrong – that is, their identity is not fixed by any causal powers it may have and it has no causal powers essentially. All we are told by Armstrong is that where an N-relation holds between two universals in a world, any instantiation of the first will be coinstantiated or followed with an instance of the second.

I showed in chapter 3 that a regularity relation between universals provides us with the same entailments as what the natural necessitation relation is *supposed* to entail; namely that all Fs are Gs. However, by the very nature of Armstrong's quidditistic properties, N, as a universal, cannot have a modal character without contravening his conception of properties. The universal, N, may play the role of a different relation in another possible world: N might, for example, play the 'taller than' relational role, where N(F,G) just means that the Fs are taller than the Gs. But Fs being taller than Gs would not be metaphysically necessary.

In our world N(F,G) is not determined by the universal generalisation – the universal generalisation, we are told, is determined by the necessitation relation. $R(F,G) \rightarrow \forall x(Fx \rightarrow Gx)$ is a *de dicto* necessity, but there is something metaphysically stronger in the case of the necessitation relation. Armstrong doesn't want N(F,G) to be *equivalent* to $\forall x(Fx \rightarrow Gx)$, and so we need some reason independent from linguistic meaning for thinking N(F,G) metaphysically necessitates $\forall x(Fx \rightarrow Gx)$.

Bird (2005) provides the Armstrongian with two options (neither of which, he concludes, helps). The first option is this:

(I*) <N(F,G)> (merely) implies <R(F,G)[58]> (Bird 2005: 150)

This material implication simply suggests a regularity between N(F,G) and R(F,G), which I would symbolise as R(N(F,G), R(F,G)). This is 'merely' a regularity, and Bird rejects it (on behalf of the Armstrongian) as a plausible candidate, for according to both Bird and Armstrong, a regularity cannot explain its instances.

[58] Bird uses R(F,G) in much the same way as I introduced in chapter 3.

Armstrong looks to be hoisted by his own petard. He claimed he could use IBE to infer N(F,G) from observed regularities, as N(F,G) is (he claims) the best explanation for these regularities; but how does he explain the connection between instances of 'N(F,G) and R(F,G)[59]' without conferring a modal character on N, a universal which, according to Armstrong's view of universals, should be categorical (if his view is to remain consistent). As we have already seen, by his own standards, he cannot explain it through a constant conjunction between the two.

Now I would of course disagree. Were there to be a constant conjunction between instances of N(F,G) and instances of R(F,G) for any F and any G, an instance of 'N(F,G) and R(F,G)' could be explained by the regularity, R(N(F,G), R(F,G)). Armstrong cannot appeal to this, and so he must postulate a further necessitation relation, which Bird characterises as the third order relation between second order relations N*(N,R). But precisely the same problem will arise when we take into consideration the non-modal character of N*. We will have to postulate a fourth order universal N**, and so forth *ad infinitum*. There is no order of N relation that (without Armstrong breaking his own rules) can ever provide the modal force Armstrong requires, and so the regress is vicious.

5.6 Concluding Remarks

The Armstrongian view is compelling to those with anti-Humean intuitions, in that it provides the metaphysical glue they desire, but ultimately the position is fundamentally flawed. The kind of necessitation the Armstrongian view proposes is 'contingent', but this succumbs to a vicious regress. To stop the regress, the N-relation must have a modal character that is constant across all possible worlds, but Armstrong's account of universals does not allow this.

[59] That is, the connection between 'the natural necessitation relation holding between two universals', and 'the regular connection between those same universals'.

Chapter 6: Scientific Essentialism

It should be becoming increasingly clear that any substantial metaphysical analysis of laws and causation will be largely concerned with the kind of properties and relations allowed in the different ontologies. To avoid the problems of quidditistic ontologies, whilst allowing for metaphysical glue in the world, many philosophers now posit dispositional properties - the identities of these properties are wholly fixed by their causal relations, and they, by their very nature, confer causal 'powers' on the things that instantiate them. Ellisian scientific essentialists embrace dispositions as the powerful properties in nature, and claim that these properties provide the oomph in our physical system – it is thus a 'dispositional essentialist' position, as at least some properties have dispositional essences. Unlike the neo-dispositionalists I will discuss in the next chapter, though, Ellis also believes there are categorical properties in the world, but that these properties are not problematic in the same way as the Humean and Armstrongian quidditistic properties are. Scientific essentialism, though, is best known for its commitment to Natural Kinds. Scientific Essentialists believe the world is highly structured by hierarchies of natural kinds with real essences, and it is this aspect of Scientific Essentialism that forms the basis of Ellis's account of laws.

In this chapter I outline scientific essentialism, its conception of laws, and the properties it commits itself to. I present arguments for and against the inclusion of categorical properties in the ontology, including a discussion of the implications of general relativity, and concluding that the neo-dispositionalist still has much work to do if he is to successfully refute Ellis's position by ruling out quidditistic properties.

6.1 An Introduction to Properties in Ellis's Account

Quidditism may not have been completely destroyed by the arguments presented in previous chapters, but nevertheless, perhaps the supposedly

counterintuitive conclusions of quidditism, combined with the failure of Armstrong's necessitation relation attempt to provide metaphysical necessity, is enough to justify a search for other options.

The proposed problems with quidditistic ontologies arise from the identity of quidditistic properties being antecedent to modal facts. The identity of dispositional properties, on the other hand, is entirely fixed by stimulus/manifestation-relations (SM-relations) central to their thesis of physical modality. Now, given that the identity of these properties is determined by their manifestations and stimulus conditions, their identity is supposed to be more than just a brute fact - it is thus metaphysically impossible for the dispositional roles they play to change in this, or any other possible world. Unlike categorical properties, then, it is claimed that dispositional properties (according to dispositional essentialists) do not possess quiddities, as their identities are fixed by their causal/nomological roles (more on this in chapter 7).

Dispositional properties are also supposed to be intrinsically 'powerful', providing the 'oomph' in causal interactions. They are instantiated in particulars, and wholly determine how these particulars behave. Dispositional properties are what might be called 'modally active', as, if there is such a thing as necessary connection between distinct existences (as they believe), the dispositional essentialists would hold that these connections exist at least partly in virtue of properties with dispositional essences.

In chapter 7 I discuss neo-dispositionalism, the view that *all* properties have a dispositional essence, but Ellis believes that the world consists of both dispositional properties, and categorical properties – the categorical properties being 'powerless' properties connected with structure and numerical relations. Ellis argues that categorical properties need not pose so much of a problem as Bird suggests, though – indeed, the properties he identifies as having quidditistic natures are some of the most 'knowable' properties of them all.

Before entering into the debate about what I, following Bird, call the 'mixed view' (that there are both categorical and dispositional properties) is tenable, it's worth looking at Ellis's scientific essentialism in more detail.

6.2 An Ontology of Natural Kinds

Ellis's (2000) ontology is one rich with natural kinds. He holds initially that there are natural kinds of processes, natural kinds of substance, and natural kinds of properties. (He later (Ellis 2010: 57-62) reduces this to just natural kinds of process and natural kinds of properties, by fitting the category of substances into the category of processes, but going into further detail here is unnecessary). Process natural kinds have dynamical members (for example: the decaying of a carbon-14 isotope), substance natural kinds have substances as members (for example: a cabon-14 isotope), and natural kinds of properties/relations have properties and relations as their members. Members of natural kinds usually have both accidental and essential properties: accidental properties can be gained and lost without that substance/process/property ceasing to be a member of the kind it was before the changes, but central to scientific essentialism is that if a substance/process/property does not instantiate all the *essential* properties of natural kind, K, then that substance/process/property is not a member of the natural kind, K. Hence if it is an essential property of an electron that it has charge -1, an electron *cannot* have charge +1. Scientific essentialism is thus what Ellis calls a scientific realist metaphysic. Natural kinds are out there in the world whether we have discovered them or not, and, according to Ellis, the very aim of science is to discover what the natural kinds are, and importantly, which properties are essential to which natural kinds.

Ellis's is an ontology of real metaphysical necessity, where 'what is metaphysically necessary is what is substantively true in all worlds in which the things or kinds of things referred to exist, and vacuously true in all possible worlds' (Ellis 2000: 335). Following Kripke (1972), Ellis

believes that since in our world we refer to H_2O as water, if we used the term 'water' to refer to any substance other than H_20, even if it looked and behaved precisely as H_2O does, we would be saying something false. But 'water is H_2O' is not *de re* necessary in virtue of the '*nominal* essence' of water, where the nominal essence of a kind is 'the set of powers or structures that a thing must have, or perhaps just the set of predicates that must be satisfied by something, for it to be called a thing of that kind' (Ellis 2009: 58), for that would only yield *de dicto* necessity. The *de re* necessity the scientific essentialist is interested in is derived from the *real* essences of things. (Henceforth when I refer to X being an essential property of *y*, X refers to a *real* essence). These real essences are 'the set of powers or structures that a thing must have for it to be a member of that kind' (*ibid*). The real essences can only be discovered by empirical investigation.

Take, for example, the law of conservation of momentum. We can imagine plenty of other possible worlds in which this law does not hold. But if this law is an essential property of the actual world (as Ellis would suppose), it is not contingent in the actual world, as, if the law of conservation of momentum did not hold, it wouldn't be the actual world. The law of conservation of momentum thus holds with *de re* necessity.

Many more questions need to be asked of this natural kind ontology, though. Here are a few that I will try to cover in this chapter:

(a) What are the natural kinds?

(b) What is the hierarchy of natural kinds?

(c) What counts as a law in a scientific essentialist ontology?

(d) What counts as a causal law/causation in a scientific essentialist ontology?

(e) Are the laws necessary for the scientific essentialist?

When we first think of natural kinds, we tend to think of biological species like eagles, cats, sea-bass and orchids; of birds, mammals, fish and flowers; of animals and plants; of biological and non-biological categories. And indeed a hierarchy does seem to emerge: each species seems to fit into at least one of the higher-order species - all members of the eagle-kind are also members of the bird-kind; all members of the bird-kind are members of the animal-kind; all members of the animal-kind are members of the biological kind, etc. But biological kinds are rarely thought to be genuine natural kinds by essentialists. There are many reasons to deny biological kinds to be genuine natural kinds. Louis Menard writes:

> We are no longer interested in the conformity of an individual to an ideal type; we are now interested in the relation of an individual to the other individuals with which it interacts. To generalize about groups of interacting individuals, we need to drop the language of types and essences, which is prescriptive (telling us what finches should be), and adopt the language of statistics and probability, which is predictive (telling us what the average finch, under specified conditions, is likely to do). (Menard 2001: 123)

Biological kinds do not, therefore, really 'carve nature at the joints' – they are merely the product of convenient linguistic conventions put in place for us to specify similarities in genetic structure, or abilities for individual members of the so-called species to mate with one-another, or whatever other meaning we decide to attribute to the term 'species' – we fix their boundaries by defining them into existence. So if biological kinds are not natural kinds, what *are*, and what are the criteria for being a natural kind?

For scientific essentialists like Ellis, natural kinds are just those categories that truly carve nature at the joints. Natural kinds are 'objective mind-independent kinds of things in nature, [and] to believe in natural kinds one must believe that things are divided naturally into categorically distinct classes' (Ellis 2009: 57). One can categorise an individual because for every natural kind, K, there is at least one property P, such that P is an

essential property of K – an individual which is a member of K can be identified as such, because it instantiates all the essential properties of K.[60]

All three natural kind hierarchies have a species-structure analogous to the biological structure considered above. There are higher-order species with subspecies. There is a highest-order species which is the subspecies of no species, and there are 'infimic' species (*ibid*) at the bottom of the hierarchies - species that have no subspecies. For Ellis, the highest-order kinds are known as global kinds: in the case of substances this is the class of all physical systems, in the case of dynamic kinds this is the class of all events/processes, and in the case of properties and relations this would presumably be the class of all properties and relations. There are therefore very few global kinds, but a huge number of infimic species. Nonetheless, every one of these species will have essential properties that uniquely pick them out from the other species.

6.3 An Introduction to Essentialist Laws

The aim of science, says the scientific essentialist, is to discover what the natural kinds are, and what the essential properties of those kinds are. According to Ellis's view of what it is to be a law of nature, this can simply be rephrased as "the aim of science is to discover the laws of nature", as for the scientific essentialist, laws describe the essential properties of natural kinds - 'all electrons have charge -1' is a law because having charge -1 is an essential property of being an electron.

> Essentialists believe that the laws of nature describe the essences of the natural kinds. This is the thesis of dispositionalism. The global laws describe the essences of the global kinds, and hence refer to all things in their respective categories; the more specific laws refer only to the more specific kinds and their various subspecies. (Ellis 2009: 64)

[60] The same could be said of an individual; that is, for any individual, x, there is at least one property P, such that P is an essential property of x. But x is not a natural kind, it is an individual.

Mumford complains that 'this would raise the question of whether such laws are available far too freely' (Mumford 2004: 133), and it is certainly true that Ellis's metaphysics of laws leads to a vast number of laws, only limited by the number of natural kinds and the number of properties essential to them. However, I do not think this should concern the scientific essentialist at all. It is perfectly acceptable, I think, for there to be many laws, so long as some of them can be deemed more fundamental than others, and that there are relatively few of the most fundamental laws. The law of conservation of momentum, then, is a global law – it is an essential property of the actual world, or at least of the natural kind the actual world is a member of. It is also a law of nature that diamonds are composed of carbon atoms, but this is a lower-order natural law. This is again in line with our intuitions.

I will talk more about Ellis's *causal* laws, but before I move on to causal laws we must first understand the rest of Ellis's ontology - in particular, the kinds of property that he allows.

6.4 Properties in Scientific Essentialism

With respect to his mixed view of properties, Ellis writes that:

> For substantive kinds,… [the] intrinsic properties or structures must include at least some causal powers, or other dispositional properties. Complex objects might have distinctive structures. Isomers, for example, may be thus distinguished. But as we descend into more elementary things, structures involving relationships between parts necessarily drop out, and, at the most elementary level, there is no structure at all. Therefore, the most elementary things existing in the world must be essentially distinguished from each other not by their structures, but by their dispositional properties alone. (Ellis 2009: 63)

But in order to describe dispositional properties we need to know the stimulus and manifestation of each disposition, and these SR-relations require reference to structure - even to 'describe the circumstances of a thing's existence' (*ibid*) we need to demonstrate the relations it stands in relative to other existents with which it might causally interact. According

to Ellis, at the fundamental level we need both dispositional and categorical properties. This is hotly disputed by neo-dispositionalists, as we will see in the next section.

6.5.1 Categorical Properties in Scientific Essentialism

Dispositional properties have their identities fixed by their causal role, but this is not the case for all properties in the Ellisian ontology. In particular he believes that both locational and structural properties must be categorical. This of course directly contradicts the neo-dispositionalist claim that *all* properties are essentially dispositional.

Ellis provides a number of arguments in favour of structural properties being categorical, but he also claims that they are quiddities (treating them as distinct). Before moving on I should make what Ellis sees to be the categorical/quiddistic distinction a little clearer. Ellis defines a categorical property as:

> a property whose identity depends on what [it is] – but not, apparently on what [it does. These are] the spatiotemporal and numerical relationships that are required to describe the structure of things (See http://philpapers.org/browse/dispositional-and-categorical-properties - Ellis *Caual Powers and Categorical Properties*: p. 4)

The categorical properties are essentially passive, 'since there is nothing that their bearers are necessarily disposed to do just in virtue of their having these properties' (*ibid*). Note that Ellis does *not* deny that structural properties play a role in laws or causation, he only denies that the properties are themselves powerful.

Quidditism is the view that properties can have their identity primitively – we have already seen this in both Lewis and Armstrong. Barker writes that generally 'a property *possesses* a quiddity just in case its identity is fixed by something independent of the causal/nomological roles it may enter into... [a categorical property] is paradigmatically... thought of as a *property* whose identity is fixed by a quiddity (Barker 2009: 1).

Ellis, though, states that all categorical properties *are* quiddities. There seems to be some conflict between Barker and Ellis here. For Barker, categorical properties have identities in virtue of their possession of quiddities. For Ellis and Bird, categorical properties *are* quiddities. For the purposes of this thesis, I use the term 'quiddity' to denote a primitive identity-fixer. But to avoid confusion as far as possible, when a property has its identity fixed by a quiddity, I shall call it a 'quidditistic property'. If all categorical properties are quidditistic, they are all properties whose identities are fixed primitively, and thus independently of their causal/nomological roles – this is what ultimately leads to the supposed problems of permuting causal roles across (and perhaps even within) possible worlds.

Dispositional monists dispute the claim that there can be properties with identities independent of their causal role, often citing examples such as how the atomic structure of a diamond gives it certain causal powers - a diamond is entirely constituted by carbon atoms, but so is graphite, and graphite clearly has very different dispositional properties from diamonds. This is because the carbon atoms in graphite are structured in lattices, whereas the atoms in diamonds have a diamond structure at the atomic level. So how might Ellis respond?

Well, he could say that even at this macroscopic level, structure does not determine the object's dispositional properties. A knife-shaped object made of butter would not have the power to cut. But dispositional monists can again take issue with this claim. They can argue that despite the knife-shaped lump of butter not being able to cut steak, it still has the power to do so. It's just that this power is 'intrinsically finked'[61]. The moment this 'knife' comes into contact with the steak it is finked and melts. I think this response deals with Ellis's initial objection.

A second objection might be that shapes are mathematical objects existing entirely independently of particulars with dispositional properties.

[61] This argument was presented by Matthew Tugby at the 2010 Metaphysics of Science conference at The University of Nottingham

The equation of a circle, for example, is $(x - h)^2 + (y - k)^2 = r^2$, where h and k are the x-coordinates and y-coordinates of the centre of the circle and r is the radius. But on behalf of the dispositional monist I would simply respond that we are not interested in abstract mathematical objects, but the properties of objects in the world. Dispositional monists do not claim that abstract objects like numbers have dispositional properties, just that all the monadic natural properties instantiated in the world are dispositional.

But even if we grant the dispositional monist that their responses to the above objections are satisfactory, Ellis believes he can argue that locational properties are quidditistic, and thus structural properties must be categorical. I will summarise these arguments as (A1) and (A2) respectively.

6.5.2 Arguments for Quidditistic Locations and Structures

A1:

1. Instances of causal powers must have locations (from where they act), and these locations must be contingent.

2. Instances of location do not have contingent locations – where something is located, it is necessarily located. (No instance of a locational property could be instantiated elsewhere)

3. Location is not a dispositional property (1,2)

4. If location is not a dispositional/powerful property, it is a categorical property.

C. Location is a categorical property (3,4).[62]

On the face of it A1 is fairly convincing. Notice though, that in the argument presented above there is no mention of the shape of particulars being categorical, only the instances of locational properties. But Ellis

[62] For Ellis's full text, see Causal Powers and Dispositional Properties, http://philsci-archive.pitt.edu/archive/00004749/

believes A1 should lead us to the conclusion that the shape/structure of a particular must be a categorical property. This requires a second argument:

A2

(a) 'The actual locations of things depend essentially on their locations relative to things whose actual locations are taken as given' (Ellis 2010: Causal Powers and Dispositional Properties, http://philsci-archive.pitt.edu/archive/00004749/)

(b) If relative locations have causal powers, then actual locations of things have causal powers (a)

(c) Actual locations do not have causal powers

(d) Relative locations do not have causal powers (b,c)

(e) The shape (or structure) of a particular is essentially the relative locations of its parts (for example, a sphere can be thought of as the three dimensional shape, all of whose locational properties lie equidistant from its centre)

(f) Shape does not have causal powers (d,e)

(g) If shape does not have causal powers then it is a categorical property

C. Shape is a categorical property (f,g)

Now let us take a look at these arguments in more detail, and see what the dispositional monist might have to say about them. The first premise of A1 looks indubitable. If I asked a dispositional monist whether this instance of the property 'being an electron' could have been located somewhere else, he would surely have to say "yes". This would be tantamount to asking "could this electron have been located somewhere else?", and the location of an electron is contingent: as Black says, 'it is clear that the same quality *can* be instantiated at spatially separated locations' (Black 2000: 94).

Similarly, *prima facie* the location of a locational property is necessary to its being the location it is. If it was not in that location it would not be that property.

Premise 3 is logically entailed by premises 1 and 2.

Premise 4 looks analytically true, as to be a categorical property is to have identity fixed independently of causal-nomological roles.

The argument is valid, and on the face of it sound, so where should the dispositionalist begin the attack? It seems to me that premises 2 and 4 are both non-trivial.

Firstly, is it really true that where a locational property is located it is located necessarily? It seems to me that it is at least *de dicto* necessary that *this* instance of a locational property is located here, as surely our conception of 'location' requires that the location of a locational property instance be fixed, but is it *de re* necessary? Paradigmatically a categorical property is a property whose identity is fixed by a quiddity. If locational properties are quidditistic, why couldn't a locational property, F, be permutable; that is, F could have been another locational property, or even the property 'ravenhood'? In accepting quidditism, on the face of it, Ellis is committed to the location of locational property instances being contingent *de re*.

However, I will argue that Ellis's ontology of natural kinds may well provide a more substantive response for him here, as he can plausibly claim the relations a location bears to other locations are essential to it; the nature of a location is fixed by its position in a network of locations, not by a network of dispositions. The proposal here is that properties can have their identities fixed non-primitively, but not by their causal/nomoligical role (and so they remain categorical properties). I will return to this possibility shortly, though, concluding that at least some primitives are still required to fix the identity of a locational property.

Secondly, dispositional monists are committed only to monadic properties having dispositional essences. ‘X is a monadic categorical property (MCP)’ is semantically entailed by ‘X is not a dispositional property (DP)’, only if the former is defined as the negation of the latter. But this is not the case. It is true that MCP→¬DP, but this is not a biconditional. A monadic categorical property is (usually thought of as) a monadic property whose identity is fixed by a quiddity, but our world is not entirely composed of monadic properties – there are also relations, dyadic properties, etc. So is location a monadic property the dispostional monist is committed to see as being powerful?

Black states, when discussing the Lewisian quidditistic ontology, that the geometrical relations are the only fundamental relations: ‘Spatiotemporal points bear certain fundamental relations to one another – geometrical relations. On Lewis’s view, these geometrical relations are probably the *only* fundamental relations, all other fundamental qualities being monadic’ (Black 2000: 91). Any location thus bears specific geometrical relations to other locations - if the relations change then the location changes, so the relations a location property bears to other locations are essential to it being the property it is. But does a location have any other essential properties? I don’t think it does. Ultimately, it seems to me that a locational property can be seen as an *n*-adic set of relations, or in other words, an ordering.

But an ordering does not automatically pick out *this* set as a location. A location is an ordering of *spatial* relations, so quiddities are still required to fix the identity of a set of relations as being a location rather than, say, a load of numbers written down on a piece of paper. Locational properties, then, are partly identified by an ordering (thus restricting the kinds of property it could be), but these relations still need to have their identities as spatial relations fixed by quiddities.

But should the dispositional essentialist be concerned? Most dispositional monists are committed only to all intrinsic monadic

properties having dispositional essence. As they 'maintain… that the nature and identity of a relation are exhausted by the relations it bears to all other relations' (Mumford 2004: 192). According to Mumford, a relation need not have a dispositional essence. If locations are sets of relations/orderings (which, for the categoricalist, must be primitively spatial relations), perhaps locations are not examples of the monadic properties the neo-dispositional must show to be dispositional[63]. The dispositionalist cannot use this response, however, not least because, as we shall see in the next chapter, his own account requires properties to be relationally constituted –the properties need to be metaphysically constituted by their (non-powerful) relations, in a similar way to how objects can be seen to be constituted by natural properties. I take it that Ellis has successfully demonstrated here that location is not a dispositional property.

So suppose Ellis has shown that locations are quiddities. What Ellis *wants* to show, is that shapes (and all other categorical properties) are quiddities – locations don't seem to be intrinsic monadic properties of things, and so the dispositional monist might not even be overly concerned about this conclusion. However, shapes look to be intrinsic, so were shape/structural properties to turn out to be categorical, the dispositional monist might be more troubled.

A2 is an attempt to move demonstrably from the conclusion of A1 (that spatial locations are categorical properties) to the substantive claim that shapes are categorical properties. This is non-trivial, even if we accept that locations are categorical. Assuming premise (a), it does seem to follow that relative locations are also quiddities. (In fact, given that I've defined location in terms of its spatial relations to other property instances, then 'relative location' and 'location' do not differ in any significant

[63] In fact, at first glance it would be rather strange if they did. The location of a particular does not add or subtract to its powers – it may well play a causal role, in so far as the spatial location of a property instance relative to other property instances will affect what stimulus conditions are met, what properties manifest, and which properties are finked; but the location itself does not change the potencies. We will see shortly, however, that Bird thinks locations, or at least spacetime points, *do* have dispositional essences.

respect. The only difference lies in the relata: with spatiotemporal locations we are talking about spatiotemporal locations to *all* property instances, but with shapes we are just talking about location of one part of an object relative to another). The move from (a) to (b), then, I deem to be an acceptable one.

Premise (c), however, is not derivable from premises (a) and (b). It stands alone, and Bird appeals to General Relativity to mount an attack. I will return to this shortly. Of course, given (c), (d) follows if (a) is acceptable, but again the neo-dispositionalist is not committed to accepting premise (e), either. The dispositional monist may well want to claim that shape, or structural properties in general, are not simply the relative locations of an object's parts - they are monadic properties that play a distinct causal/nomological role. And, given the numerous examples they can provide, it is easy to see why – even as children, we learn to use the spade over the bucket to dig holes, for it seems to have the power to shift sand! The dispositionalist might then reject premise 3 on two counts: firstly that location does play a dispositional role, and secondly that at least some dispositional property instances have their locations necessarily; namely, instances of location!

6.5.3 General Theory of Relativity (GR) and Structural Properties

Just using medium-sized objects like spades to assert that structural properties have dispositional essences is not likely to persuade Ellis. After all, Ellis does not deny that structural properties and spatial relations play a role in laws and causal interactions (if he did, his position would be flagrantly flawed). The dispositional monists need something more than what is essentially an appeal to intuition, and Alexander Bird argues that, when we consider the implications of GR (combined with his assumption that we should adopt the contemporary view that interpretations of GR should be background independent), the spatiotemporal locations Ellis sees as quiddities are either not part of our ontology at all (and hence the

dispositional monist need not worry about them), or they are spatiotemporal points with dispositional essences.

In chapter 8 I will evaluate the various metaphysical conceptions by looking at them in the context of the Principle of Least Action. I argue that any serious contender must be compatible with it, and preferably explain its holding. I think compatibility with least action principles is particularly important, as it has fallen out of *all* our serious attempts at mapping the evolution of the system (Newtoniam Mechanics, Special Relativity, General Relativity, String Theory etc). It seems to me that if there are any *a posteriori* necessary conditions of being a law, adherence to this principle is likely to be one of them. GR is not on this scale of generality – just as Newtonian Mechanics was inconsistent with GR, so it is possible that GR will be incompatible with what is arguably the ultimate aim of physics: a unified theory of quantum gravity. However, given its wide scope and applicability, and its consistency with empirical evidence, it seems to me we have good reason to take the precepts of GR seriously.

The two main features of GR important to this discussion are (a) that although simultaneity was frame-of-reference dependent under Newtonian Mechanics and Special Relativity, in GR the locations in *spacetime* can be agreed upon from all frames of reference. This gives rise to Einstein's principle of general covariance, and allows the laws of nature to be invariant under arbitrary co-ordinate transformations. We should, from a metaphysical perspective, cease to take seriously the independence of space and time when considering the physical system in its entirety (when considering the 'world-line' of a particular we can, in a sense, allow time to be measured by the watch of that particular and observe its changes, but the laws of nature are not localised in this way); and (b) it accounts for empirical observations incompatible with previous physical theories by employing the *curvature* of spacetime[64]. In developing a

[64] The equivalence of acceleration and gravity central to relativity, and the curvature of spacetime that came out of it, was demonstrated in 1919 when the trajectory of light from a distant star must have bent in order for it to be observed. Given that photons have no mass, this

metaphysics of location in spacetime, we must also bear in mind the implications of spacetime *being* curved, what effects the curvature of spacetime might have on the things *in* spacetime, and what *makes* it curve.

Bird tries to take advantage of these considerations, and argues that GR gives us good reason to think that spacetime points have dispositional essences. Unlike the Newtonian view, in which the physical system evolves against a background of space and time unaffected by the laws of nature, contemporary physicists increasingly endorse background-free ontologies (this is probably not the orthodoxy, yet, but Bird pushes this issue so I'm happy to follow him, here) - where background-free ontologies either:

1. remove space and time altogether, in which case spatiotemporal relations fall out the picture entirely – this is known as relationalism about spacetime[65] - in denying the ontological priority of spatiotemporal points, we are denying that background relations are the base; or

2. they include space and time in their ontology, but not as a background. There are spacetime points as real existent entities in our world, but one can have the same fundamental spacetime with different spatiotemporal points - this is a background-free version of substantivalism about spacetime.

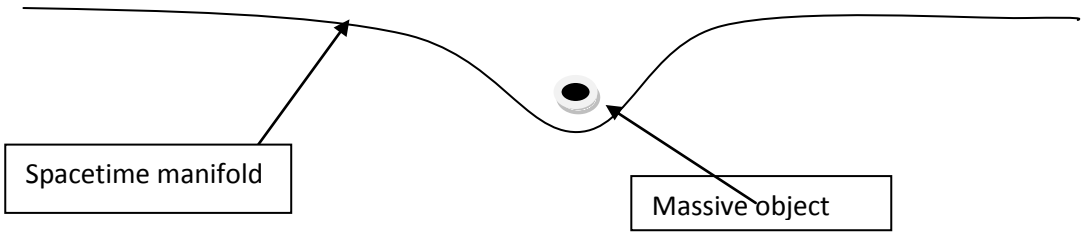
If we endorse (1) then, according to Bird, the dispositionalist has nothing to worry about, as the dispositional monist does not claim that non-fundamental properties are dispositional. If we endorse (2), Bird argues, the dispositional essentialist *is* committed to a dispositional account of locations, but this is not problematic because under a substantivalist account of spacetime, spacetime does indeed seem to have a dispositional essence.

---

was possible only if spacetime is curved – so we certainly have good empirical evidence in favour of spacetime-curvature

[65] As a matter of fact, Vassilios Livanios claims (1) is a '*radical* interpretation of background-independence' (Livanios 2009: 388)

According to standard interpretations, spacetime in GR is not, as in the Newtonian picture, inert and unaffected by the things that occupy it, but a manifold with a structure subject to change in virtue of the movements of massive objects within it. This is often illustrated with the analogy of bowling ball (which represents a massive object) on a trampoline (which represents the spacetime manifold), where weight of the ball curves the trampoline's surface.



If the principle that matter *tells* spacetime how to curve is right; that is, there is a causal relation between the distribution of matter and the curvature of spacetime, then the idea that spatiotemporal points have dispositional properties is not implausible. As Bird puts it:

> One reason why it is difficult to see space and time as causes on a classic substantivalist conception, is that it is difficult to see them as any way being effects. The background is unchanging. But if it is unchanging how can it generate any effects [that is, have the essential quality of a dispositional property]? On the other hand, if it is subject to change, then it is easier to see how it might itself be a cause of change… In dispositional essentialist terms, we can see that by being potential manifestations of dispositional essences, spatial and temporal properties may also have dispositional essences themselves (Bird in Handfield 2009: 240).

Here, Bird is appealing to the action-reaction principle to suggest that if the location of massive objects curves spacetime, then the curvature of spacetime has a causal effect on the massive objects. Indeed, the thought that the structure of the spacetime manifold determines how matter moves (as well as matter telling spacetime how to curve) is well entrenched in

interpretations of GR: it is undeniable that in GR, test particles in inertial states follow non-linear geodesics due to the curvature of spacetime ('due' should not be interpreted here as necessarily having has causal implications, but it is easiest to use this term in here). In other words, world-lines of test particles in inertial states are not straight lines as we imagine them in classical spacetime, as the test particle follows the contours of the curved spacetime manifold – just as if we dropped a marble on the trampoline, it would roll down to the bowling ball following the shape of the elastic surface.

It looks, under this hugely simplified version of GR, as if the locations and masses of the objects in a system determine the 'curvature of the spacetime, that is to change the dynamical properties of spacetime points' (*ibid*), and the shape of the spacetime manifold determines the locations of the objects. To put it in the causal terms of dispositional essentialists:

(1) the structure or geometry of spacetime (at least partially) causally determines which paths objects follow, so the structure of spacetime is powerful; and

(2) the mass and spatiotemporal location of the objects causally determines the structure of spacetime, so the masses have the power to curve spacetime.

Structure can thus be seen as having a dispositional essence.

So what are the implications here for Ellis? Well, Ellis argues that you can take away all the powers from a location (and we must be talking about spatiotemporal points, now), and the location remains. According to the account above, the spacetime manifold is (a) affected/changed by other property instances, and (b) can act as stimulus conditions for other property instances. But are Ellis's conclusions and the account above incompatible? Before making any conclusions here, I would first like to

turn my attention briefly to the relationalist/substantivalist interpretations of spacetime, as this will affect our conclusions.

There is still a debate between the substantivalists and relationalists about spacetime (as we have already seen, the dispositionalist claims he has no issues if relationalism is true). Although I will not make a contribution to the debate here, I would like to note that the outcome is of some interest when discussing the possibility of quiddistism about location. Curvature can be accepted by both parties, as is it as intrinsic property of spacetime geometry, independent of how we describe the geometry (whether there are spacetime points, the nature of these spacetime points etc). The substantivalist, though, contrary to the relationalist, endorses the view that spacetime geometry is composed of spacetime points whose identities are not uniquely determined by the theory, where most substantivalists accept that different spacetime manifolds (at different possible worlds) *could* differ *only* in virtue of the identity of the spacetime points that constitute it.

Belot and Earman (1999, 2001) take *relationalism* about spacetime to be a denial of a fundamental ontological role for spacetime points – they 'deny that there could be two possible worlds with the same geometry which differ only in virtue of the way that geometry is shared out over existent spacetime points' (Belot and Earman 2001: 18). If we accept substantivalism, however, whereby spacetime points have their identity fixed independently of their position in the geometrical structure, there is no reason why they could not swap their locations whilst retaining the geometrical 'shape' of spacetime. When physicists have been asked whether there can be two worlds which differ solely over a permutation of the spacetime points, they answer that 'there is nothing anti-substantival about denying that there can be such distinct possible worlds' (see Butterfield,1989; Brighouse, 1994; Rynasiewicz, 1994; Hoefer, 1996). (Pooley in Rickles, French, Saatsi 2005: 107).

Again, for these substantivalists, the spacetime points are *not* identified by the geometry – they have their identity independently of the geometry. Spacetime points, for the substantivalist, are haecceistic.

So far I have been talking about the identity and properties of spacetime points, not of locations. Locations, it seems to me, are *properties* of spacetime points – properties that could be entirely distinct from the dynamical properties of the haecceistic spacetime points, and not properties essential to their being the individual spacetime points they are. Ultimately, when Bird says that (paraphrasing) "mass changes the curvature of spacetime, or in other words, changes the dynamical properties of spacetime points", it might be less misleading to say that "mass changes what locational properties the spacetime points instantiate". But even that is strictly inaccurate, as in a four-dimensional spacetime manifold, there is no *change* of this nature isn't really a part of the picture. Rather, mass *affects* the shape of the manifold. Although every geometrical location is occupied by a spacetime point, the geometry is in some sense independent of these points – as permutation of the spacetime points is a possibility. Ellis's claim, then, that you can remove all the powers from a location (the thing with the powers is located there contingently) and the location remain, could still be true; that is, if, as looks to be implied by the haecceistic nature of spacetime points, we can take the current location of a spacetime point *p*1, to have the powers located there contingently (the powers instantiated by spacetime points *p*1 and *p*2 in world, *w*, might be different[66], but the nature of spacetime allows for the permutation of p1 and p2), but the current *location* of p1 is necessarily *located* where it is (that is, *that* locational property can only be instantiated by a spacetime point at that geometrical point of the manifold), then Ellis's argument for location being a categorical property is not thwarted by Bird's response.

[66] Perhaps not in our world, but there is nothing metaphysically impossible about distinct spacetime points instantiating different causal powers even under a dispositional essentialist metaphysic, and spacetime points are particulars, not properties.

If we accept Ellis's interpretation of shape, it also follows that shape-properties are also categorical properties. The shape of an object is the relative locations of its parts; that is not to say that it is the relations between the spatiotemporal points the object happens to occupy, as you could swap all those spatiotemporal points for other ones and keep the object in the same location - it is the relative *locations* of the object's parts in the *geometry* of spacetime that constitute its shape. Shape, then, can be a categorical property with a partly quidditistic nature.

Other objections to Bird's argument have also been raised. The motion of massive objects is precisely described by Einstein's equations, but there is more than one way of interpreting these equations. Vassilios Livanios looks at the way Einstein's equations can be interpreted, and points out two notable possibilities. The first possibility, and presumably the one Bird accepts, is that 'the curvature of spacetime depends *causally* on the distribution of matter fields across it' (Livanios 2008: 389). But, says Livanios, talk of the distribution of matter presupposes a pre-existing metrical structure, which cannot be. So we need an alternative interpretation whereby 'the dynamical structure of spacetime allows the trans-world variation of spacetime structure to vary in such a way that the Einstein equations hold' (Livanios 2008: 389) We should look upon the Einstein equations as showing us what can only be described as a law-like constraint (which, as we have seen many times, does *not* require causal connections) between spacetime structure and the distribution of the massive objects in any metaphysically possible world. But, as Livanios says:

> If matter does not *cause* any changes in spacetime structure, then spacetime (better, the metric field) is not a recipient of change and, consequently, does not have any passive causal essence. Moreover, given that in this case we do not have any causal influence, the appeal to the action-reaction principle is unjustified. So, pace Bird, there is no reason (related to the application of this principle) to suppose that spacetime (metric) has an active dispositional essence. (*ibid*)

Bird's argument is beginning to look very unappealing. I have shown that the property dualist may well be able to hold on to his view that locations and shapes are categorical properties, *even if* we accept that spacetime points have powerful properties. And Livanios has given us good reason to think that matter does not cause any changes in spacetime structure, anyway.

This does not, if Bird is even half right, completely undermine neo-dispositionalism. Bird says that if we accept relationalism about spacetime then there's no need to accept categorical properties as they eliminate space and time from their ontology. He goes on to argue that the substantivalist can give a dispositional account of spacetime, but even if this fails, he can at least fall back on relationalism.

The obvious question, then, is what entitles Bird to thinking shapes are dispositional? Ellis has given us an account of shape in terms of relative locations, but Bird has not given us any such account. In fact, Bird says the dispositional monist can get out of trouble by denying that structure is a fundamental property:

> If [structural properties] are not fundamental properties, then having dispositional essences or not does not distinctly bear on the truth of dispositional monism. (Bird in Handfield 2009: 232)

By giving the dipositional monist this option, he denies he even *has* to provide a dispositional account of structure, as he is committed only to fundamental properties having dispositional essences, and structural properties might not be fundamental.

So what should we make of all this? Ellis has given us a categorical account of locations and shapes. Bird has responded by appealing to the nature of spacetime, but his attempt fails unless we adopt a relationalist account of spacetime. Bird can fall back on shape/structural properties not being fundamental, but if this discussion of General Relativity has shown us anything, it is the fundamental role that structural properties play in our world. In light of our understanding of spacetime

curvature, it seems to me that the claim that structural properties are not fundamental requires much work, as structural properties certainly seem to be fundamental, intrinsic, monadic properties of the spacetime manifold.

6.6 Dispositional Properties in Scientific Essentialism

I have outlined Ellis's ontology of natural kinds, briefly described how it involves both categorical and dispositional properties, and provided arguments for and against the requirement of categorical properties. I have, however, said little about the dispositional properties in Ellis's ontology which are so central to his view of causal processes and causal laws.

According to Ellis every natural kind of property has a real essence, and in the case of dispositional properties these essences are (of course) dispositional. The dispositional essence of 'fragility' (if it is a natural kind of property) would be, according to Ellis, linked with the manifestation and stimulus conditions described above. If a property does not have this essence, then when we refer to it as 'fragility' we would be saying something false. To illustrate this perhaps a little more clearly, imagine in our world one of the (essential) potential stimulus conditions for the property 'fragility' is 'being hit with a hammer'. Now suppose in another possible world, *w**, there is a property that can be instantiated by particulars that has precisely the same stimulus conditions and manifestations as 'fragility' in our world, except 'being hit with a hammer' is not one of the stimulus conditions. Now the inhabitants of *w** may never have tried to hit a vase with a hammer, but the truth about whether it would smash or not when hit with a hammer is out there in the world for them to discover (if they tried, it wouldn't smash). Any object with this other-worldly property, despite their having behaved in exactly the same way as all fragile objects in our world would have done (given the stimulus condition 'being hit with a hammer' have never been met), the vase would not instantiate one of the essential properties of the property 'fragility' we refer to. If we went to this world and said "this vase

is fragile", we would be saying something false – we would be referring to the property 'fragility', whereas the vase in this other world instantiates the property 'fragility*'. Similarly, if inhabitants of this other possible world referred to what we might term fragility* as 'fragility', then if they attributed the property 'fragility' to an object in our world (that would smash when hit with a hammer) they would be saying something false, as their term 'fragility' refers to the properties with all the essences of fragility*, a property whose stimulus conditions does not include being hit with a hammer.

It may look as though Ellis is really just defining 'fragility' (to use our example) as that property which displays smashing when hit, but again it must be emphasised that Ellis's intention is not to define these properties into existence. 'Fragility', if it is a real natural kind of property, is a metaphysically real property that must be discovered *a posteriori*. In science we are not discussing the *meaning* of dispositional term, but discovering the real essences of properties in the world, 'independently of our systems of classification'(*ibid*).

Ellis thus avoids both the 'swapping dispositional roles' and 'multiple realisability' objections raised against the Humeans and Armstrongians, as any property that does not have all the essential stimulus conditions and manifestations of property P (which has had its referent fixed) cannot be property P.

It is worth discussing exactly how Ellis sees dispositional properties, and to do this one has to start by looking at natural kinds of processes. These are displays of dispositional properties, the essences of which are the dispositional properties. 'In the case of any simple causal process, the real essence will be a dispositional property, and the scientific problem will be to specify precisely what this property is' (Ellis 2000: 333). In the simple non-multitrack cases (where a disposition only has one manifestation), the essence of a dispositional property can, according to Ellis, 'be uniquely characterized by an ordered pair <C,E>. Ellis defines

‘C’ as a kind of circumstance, and ‘E’ as a ‘kind of event’ (to stay in line with the general dispositionalist talk, we can think of these as the stimulus conditions and manifestation of the property, but it is important to bear in mind that both the stimulus conditions and the manifestation of the property must be thought of as natural kinds in their own right. It seems to me that Ellis has this stimulus-manifestation relation in mind when he writes that ‘if x is an object with this dispositional property, then x may be said to have the power or propensity to E in circumstances C’ (*ibid*)).

The natural kinds of process that ultimately constitute the evolution of physical systems thus occur in virtue of the natural kinds of (dispositional) properties of things, and the identity of the properties determining the evolution of the physical system are fixed by the natural kinds of process to which they refer. For any two worlds with the same kinds of properties and relations, the same kinds of natural process will occur given the same circumstances: given that causal processes are (at the very least) a species of natural processes, and that the essential properties of the natural kinds of process are dispositional properties, if two worlds have the same dispositional properties, the same kinds of causal process will occur in both worlds.

To illustrate, consider the substance natural kind ‘electron’, the process natural kind ‘subjected to repulsion force’, and the property natural kind <within negatively charged electromagnetic field (C), subjected to repulsion force (E)>. We know the following:

1. It is of the essence of an electron that it is negatively charged, so anything that is not negatively charged is not an electron;

2. The dispositional property <C,E> is an essential property of negatively charged objects.

3. If the stimulus conditions, C, are met, the natural kind of process ‘subjected to repulsion force’ occurs, as this is of the essence of <C,E>.

In all metaphysically possible worlds, then, when an electron is in a negatively charged field it is subjected to the repulsion force. Metaphysical necessities like that discussed above hold in virtue of the natural kinds of things that exist, and the properties which determine their memberships of these kinds. Consequently, the desideratum outlined previously that the causal necessities be *de re*, is satisfied.

I have shown how Ellis believes his essentialism allows him to escape the problems posed by quidditism, and how this essentialism forms the foundation of Scientific Essentialism. I must now consider the account of causal laws that falls out of this framework in more detail.

6.7 Causal Laws in Scientific Essentialism

Causal laws are, in Ellis's ontology, unsurprisingly tied up with dispositional properties as well as natural kinds, as although laws of nature are the essential properties of natural kinds, the causal processes involve the manifestation of dispositional properties. Crudely, Ellis makes the following moves to get to a formalised account of causal laws:

'Let D be the causal disposition <C,E>, and D(x,*t*) the proposition that x has this disposition at *t*.' (Ellis 2001: 130)

According to this state of affairs, when C occurs, E occurs (and E is caused by C).

'Let C(x,*t*) be the proposition that an event or state of affairs of the kind C exists or occurs to x at *t*, and E(x,$t + \delta$) the proposition that an event of kind E occurs to x in the time interval from *t* to $t + \delta t$.' (*ibid*)

If an event of type E *must* occur during the period between *t* and $t + \delta$, 'then the disposition <C,E> of x at *t* is said to be *causally determinate.*' (*ibid*)

Where this is true of *all* instantiations of <C,E>, Ellis calls this a 'deterministic law of action of the disposition'. This, he says, is a causal law, and can be formalised as:

(CL) $\forall x,t\ [C(x,t) \Rightarrow E(x,t)]$

Causal laws are thus defined both in terms both of dispositions and of natural kinds; firstly a disposition is, for Ellis, characterised by an ordered pair of natural kinds of events, and secondly the law statements (as expressed above) are universally quantified statements concerning the relations between events of *kind* C and events of *kind* E. Note that this formalism bears strong resemblance to the form of the Humean law-statement, but according to Ellis, this statement is true *in virtue* of the dispositional property <C,E> instantiated by x at *t*, and so law-statements refer to necessary connections between distinct existences, not *just* regularities.

This universally quantified statement is the form of a causal law-statement, but it must follow from the Scientific Essentialist metaphysic that For Ellis, all laws are essential properties of natural kinds. Fundamental causal laws, just like other fundamental laws, are essential properties of the world kind.

6.8 Conclusions to Chapter 6

In this chapter I have discussed the main aspects of, and problems with, Ellis's scientific essentialism. Summarising: Ellis presents a view of laws whereby laws *describe* the essential properties of natural kinds. Laws are therefore propositions, not, as with Armstrong, aspects of nature. However, these propositions do describe *de re* necessities. It is no accident that electrons have charge -1, because charge -1 is an essential property of being an electron. Thus the 'accidental regularities' objection holds no weight, and the problems of Armstrongianism seem to be overcome. Ellis's laws are also *propositions* describing the nature of the world, rather than the laws being *in* nature, in the Armstrongian sense - I find this view compelling, as it seems to me that most people, philosophers and non-philosophers alike, can meaningfully talk about laws of nature, without having any metaphysical understanding of what laws might be at their core.

Ellis believes that a complete ontology has both dispositional *and* categorical properties, where the former are powerful properties/propenstities with directedness towards their manifestations, and the latter are structural and numerical relations that, although they play a role in cause and law, are not themselves powerful properties. Ellis provides arguments for locations and structures being quiddities, and Bird responds by appealing to GR. I have shown that the arguments from GR are found wanting, and Bird's 'back-up plan' of concluding that structures are not really fundamental properties anyway, and hence should not concern the dispositional monist, is dubious. If consideration of GR teaches us anything, it is that geometrical structure *is* fundamental to the evolution of our system. It seems to me, then, that Bird has a lot more work to do if he is to rule out the 'mixed view'; that is, the view that there are both fundamental dispositional and fundamental categorical properties.

Ellis's Scientific Essentialism looks compelling on the evidence we have seen so far. However, its plausibility in part relies on the success of an account of dispositional properties. Further discussion of dispositional properties will be provided in the following chapters.

Chapter 7: Neo-Dispositionalism

7.1 Neo-Dispositionalism

In this chapter I examine the dispositionalist account of causation as advocated by Alexander Bird and Stephen Mumford, compare and contrast it with the scientific essentialism of Brian Ellis, and show how their position can be defended against a number of regress objections.

I call the view looked at in this chapter neo-dispositionalism, to distinguish it from the view of properties Bird calls 'dispositional monism' (Bird 2007: 3), (elsewhere the dispositional monist view is given the title 'pan-dispositionalism' (Bostock 2008)). Dispositional monists believe that all properties have a dispositional essence, but not all dispositional monists believe all properties are wholly dispositional. John Heil (2003), for example, holds a view in which all properties have dispositional essences, but all properties also have a categorical aspect to them. Bird's dispositional monism, however, rids itself of categoricalism altogether – not only does he endorse 'dispositional monism' (that all properties have a dispositional essence) (Bird 2007: 3), but he believes all natural properties to be *wholly* powerful/dispositional[67].

7.2 The Modal Character of Neo-Dispositionalism

It is important first to grasp the modal character of this account of causation:

Bird writes that 'in all possible worlds, any object that possesses P is disposed to yield M in response to S' (Bird 2007: 45), formalised as

(DE_p): $\Box(Px \rightarrow D_{(S,M)}x)$

From (DE_p), and the conditional analysis of dispositions as a necessary equivalence ($\Box(D_{(S,M)}x \leftrightarrow Sx \Box\rightarrow Mx)$), we can derive

[67] Note that in this discussion we are only talking about natural properties.

(I) □(Px → Sx □→ Mx).

According to neo-dispositionalists, then, in all possible worlds, if *x* instantiates dispositional property, P (which has stimulus conditions S and Manifestation M), if the stimulus conditions were met, then P would manifest.

Furthermore, when we 'consider any world *w* and any case where some x in *w* possesses the potency P, where *x* acquires the stimulus S (that is: (Px & Sx)), we can derive a universal generalisation of the form:

∀x((Px&Sx) → Mx)[68].

This is a universal generalisation of the form we often associate law of nature, and indeed Bird (but not Mumford) believes a theory of laws as supervening on dispositions can be developed. I will provide a more detailed discussion of this possibility in the second half of this chapter.

7.3 Dispositional Monism v. Ellisian Dispositionalism

Looking at the preceding sections, one can identify the main differences between the neo-dispositionalist and the Ellisian metaphysics:

Firstly, the former deny that there are any categorical properties, avoiding the supposed problems posed by their quidditistic nature. Its qualitatively more parsimonious ontology might be seen as an advantage over the mixed view;

Secondly, the neo-dispositionalists do not require an ontology of natural kinds (note, though, that there being natural kinds with real essences is not incompatible with dispositionalism);

Thirdly, the modal character of Bird's dispositional properties provides the neo-dispositionalist with the opportunity to develop an account of laws '*in* nature', supervening on this one kind of property, whereas For Ellis, laws are propositions describing the essential properties of natural kinds.

[68] From (I) and (II) we have (III) Mx, discharging (II) we have (IV) (Px&Sx) → Mx, and since *x* is arbitrary we may generalize ∀x ((Px&Sx) →Mx) (Bird 2007: 46)

The purpose of this chapter is to determine whether or not neo-dispositionalism is a plausible account of properties, and to see out how one should view dispositions. In this next section I look at two questions that arise: that of whether properties are clusters of powers, and that of whether or not there can be multi-track dispositions.

Once I have outlined the details of the most plausible dispositional essentialist ontology, I move on to some pressing arguments against the position, namely the regress objections.

7.4 Properties, Powers, and Multi-Track Dispositions

Stephen Mumford (2004) puts forward the idea that (at least some) properties might be 'clusters of powers' – that is to say, a single property might consist of a multitude of powers with distinct stimuli and manifestations. Elasticity, for example, might be considered a single property, yet it has the power to bend, to stretch, to bounce, and so on. But elasticity is not one of the fundamental properties I'm interested in, in this thesis. It is no doubt true that we tend to group powers together such that when we attribute a property to a particular, we are attributing to that object a number of powers that regularly come as a package. The power to bend, to stretch and to bounce, often come together, and so for the sake of convenience we give a name to this group of powers. But it seems to me that these powers can often exist independently of one another.

To justify his claim, though, it might be argued that Mumford only needs to show that *one* property has a number of powers that cannot come apart; that is, there is a fundamental property such that one of its powers cannot be instantiated unless the power or powers associated with it is also instantiated. The clusters of powers view can has been repackaged in the form of 'multi-track dispositions', so I will focus on this discussion.

Multi-track dispositions take a number of different forms: they are either:

(a) dispositions with one stimulus condition but multiple possible manifestations (disposition D has stimulus condition S1 and manifestation M1 v M2 v M3…);

(b) dispositions with one manifestation but many possible stimulus conditions (disposition D has stimulus conditions S1 v S2 v S3…, and manifestation M1);

(c) dispositions with more than one possible stimulus condition and more than one possible manifestation (disposition D has stimulus conditions S1 v S2 v S3… and more than one possible manifestation M1 v M2 v M3…)[69] ; or

(d) dispositions with either a conjunction of stimulus conditions, and/or a conjunction of manifestations (disposition D has stimulus conditions S1^S2^S3... and manifestation M1^M2^M3...) (see Bird 2007: 21-24).

If bending and bouncing are different manifestations of elasticity (an elastic object having the power both to bend and to bounce under different stimulus conditions), for example, then elasticity would count as a multi-track disposition. But, it seems to me, the stimulus conditions and manifestations associated with elasticity can be broken up into a conjunction of multiple distinct single-track dispositions. If these single-track dispositions constitute the multi-track disposition, then the latter cannot be fundamental. Consider the case of the stimulus condition 'massive object being thrown at window', and its manifestations 'window smashing', and 'making a noise'. One might be tempted to see mass as giving things that instantiate it a multi-track disposition, as in virtue of the object's instantiating this property, throwing it at a window causes it both to smash a window and make a noise. But these can be broken down into two separate, more fundamental dispositions: the disposition to smash a window when thrown at it, and the disposition to make a noise when thrown at a window. There is no need to posit multi-track dispositions.

[69] Note that there these disjunctions do not entail that any one of S1, S2, S3 could produce any one of M1, M2, M3. It may be the case that S1 can only produce M1, S2 can only produce M2 and S3 can only produce M3.

In cases of multi-track dispositions of type (a) the property's fundamental constituents all share the same stimulus condition, but they have distinct identities fixed by their different manifestations. With those of type (b) the multi-track disposition is constituted by more fundamental properties, whose identities differ in virtue of the different stimulus conditions (but having the same manifestation). With those of type (c) the constituent dispositions may be even more obviously distinct: the multi-track disposition could be constituted by distinct properties with both different stimulus conditions *and* different manifestations. Now consider dispositions of type (d): suppose that for water to shatter, the temperature of the water has to be sub-zero degrees centigrade, and it must be hit with a hammer – these are distinct events, whereby water can be frozen without being hit with a hammer, and vice versa. This seems to be a case of genuine multi-track disposition, but it is not a fundamental one. Let us look at few examples of the more fundamental dispositions:

(1) The disposition to attract massive objects

(2) The disposition to attract protons

The first notable aspect of the more fundamental dispositions is that they tend to be functional. They tend involve fields of one kind or another, and the strength of those fields depends on certain circumstances. An object instantiates disposition (1) if it, itself, instantiates the property 'mass'. But the attractive force it exerts on other massive objects is dependent on how 'massive' the object is. One might be tempted to interpret this disposition as being multi-track, as the disposition to attract massive objects seems to have multiple stimulus conditions (the different quantity of mass), and multiple manifestations (the different strengths of the gravitational fields emitted). But it seems to me that there is no need to allow for multi-track dispositions, here. The gravitational force between two objects is directly proportional to the two masses, and inversely proportional to the square of the distance between them. Each value of mass contributes towards the

attractive force in a different way; that is, specific quantities of mass give objects different dispositions.

Ultimately, if we want the properties instantiated by particulars to tell us exactly what powers that particular has, knowing that the particular has mass, or that it has charge is insufficient. We need to know the quantities of mass and charge, as each quantity confers different dispositions.

7.5 The Regress Objection and Relational Constitution .

Although neo-dispositionalism does seem initially to have significant virtues over its mixed-view and categorical monism opposition, there are a number of criticisms that need to be dealt with. The first objections to be tackled that are specific to the dispositional monist positions[70] (as opposed to ontologies allowing for dispositional properties *as well* as categorical properties, or categorical monsm) are the 'Regress Objections', which come in three different forms – I shall call these the 'not enough reality regress' ((Armstrong 1983, 2005; Blackburn 1990)), the 'epistemic regress' (Swinburne 1980), and the 'regress of identity' (Robinson 1982).

7.5.1 The Ontological, or 'Not Enough Reality' Regress

The first criticism is that *if* a dispositional property is just a power to produce some other power, ('the power to produce A is nothing but the power to produce the power to produce B… and so on' (Armstrong 1983: 123)), and all properties are dispositional, then there is no 'being' in the world. Or to put it another way: if both the stimulus and manifestations of a disposition are also just potencies, then there can be no 'action'.

The 'not enough reality' regress suggests that there needs to be more than just powers in our ontology, or particulars are just constantly changing their potencies. Potencies are never manifested, they are just replaced by other potencies, and replacement by other potencies is not a

[70] These objections are raised against dispositional *monist* positions (including neo—dispositionalism)

genuine manifestation. It is thus argued that an ontology of pure powers is untenable. This objection does not, you will note, itself rule out the possibility of dispositional properties. Its upshot is that there must be *at least some* categorical properties in a workable ontology.

Unsurprisingly the neo-dispositionalists argue that this regress fails. The argument is that powers do not have enough actuality; that is, that a change in potency does not count as an actual manifestation, and so the regress works only if powers are nothing in themselves – that is to say, the changing of a particular's potencies does not constitute a genuine manifestation as we normally see it. A genuine manifestation, I take it, must be the gaining or losing of natural, first-order properties: properties like colour, shape, charge etc. These properties are not, or so the objector would hold, potencies. However for the neo-dispositionalist, the properties listed above *are* dispositions. The objector is just unjustifiably assuming that gaining charge is not (merely) a change in dispositions.

Of course the claim that dispositions are genuine entities in their own right (not reducible to counterfactual conditionals) requires argument, but if the categoricalist raises this issue, the dispositionalist can respond that the claim that categorical properties have sufficient reality, and dispositional properties do not, cannot be right, as the essential features of Lewis's Humean categorical properties are all also features of dispositional properties. Bird outlines the features of a categorical property thus:

(a) It is distinct from (not identical with) other properties;

(b) It is a universal and thus can have instances;

(c) For some *n* it is an *n*-adic universal. (Bird 2007a: 11)

All of criteria (a), (b) and (c) are, Bird claims, also true of dispositions. Dispositions just have, as an extra criterion, a dispositional essence.

However, it seems to me that Lewis could, and indeed would add the following to his list of essential features of categorical properties:

(d) Properties are sets of particulars (if (d), then (b) must be false, and vice versa)

This nominalist position grounds properties in particulars, so it cannot be argued that Lewis's properties do not have enough 'being'. The nominalist's properties are abstract entities in the form of sets, but the members of those sets are particulars occupying spacetime points – the particulars play a constitutive role, giving the property its 'reality'. Alexander Bird's properties do not share this feature (the property is supposedly wholly located wherever it is instantiated, but these instantiations play no constitutive role with respect to the universal). If it is this feature that gives properties their reality, then Bird's argument that the categoricalist fares no better, fails against the nominalist versions. However, the categoricalists who hold a view of properties as universals (Armstrong, Dretske, Tooley amongst others) cannot use this argument against Bird.

Armstrong (2005) re-packages the not enough reality objection when considering the view that causation is the transference of potencies. 'Causality becomes the mere passing around of powers from particulars to further particulars' (Armstrong 2005: 314). Mumford quickly dismisses this objection, though, simply pointing out that causality's being the passing around of powers is actually a very attractive account (see Mumford 'Passing Powers Around', forthcoming). In certain circumstances I can see how Mumford's claim here is plausible. I think the closest the natural sciences get to (explicitly) using potencies in physics, is in the use of 'potential energy'. A coiled up spring has elastic potential energy. When the spring is released, the ball that was sat on top of it shoots into the air gaining gravitational potential energy, and so on. It is not, therefore, on the face of it at all obvious that the gaining and losing of potencies should not be considered genuine manifestations. I conclude, then, that this particular regress objection is unsuccessful.

7.5.2 The Epistemic Regress

The epistemic regress (Swinburne 1980), as the name suggests, concerns our knowledge of dispositional properties. If all there are are potentialities then we can never know what properties anything has, and yet it is obvious that we do! We can only attribute a property to an object if we have seen its effect (another property), but we can only know the 'effect' property has been instantiated if it, in turn, manifests, and so on…

Bostock (2008) provides a similar response to the epistemic regress as he does to the ontological regress; namely that it only works if we presuppose that properties are not entities in their own right. According to Bostock we can know dispositional properties directly, so knowing the properties that caused them is not necessary. He writes that 'Pan-Dispositionalism claims that properties, of their nature, are caused to be instantiated by certain other properties and, in turn, cause, in certain circumstances, other properties to be instantiated. But this metaphysical claim in itself says nothing about how we become acquainted with properties' (Bostock 2008: 147). Again, I don't think the epistemic regress is overly problematic for the neo-dispositionalist.

7.5.3 The Identity Regress

The identity of a categorical property is fixed intrinsically; its identity is primitive[71].

In contrast, the identities of dispositional properties (for the neo-dispositionalist) are fixed *relationally*; that is, the first order natural properties have their identities fixed by the relations they stand in to their stimuli and manifestations. These stimuli and manifestations are natural properties in their own right, and so these, too, have their identities fixed by the stimulus-manifestation relations (SM-relations) they stand in to other natural properties. This seems to require no appeal to any primitives of the nature required by the categoricalist. On the face of it there is a

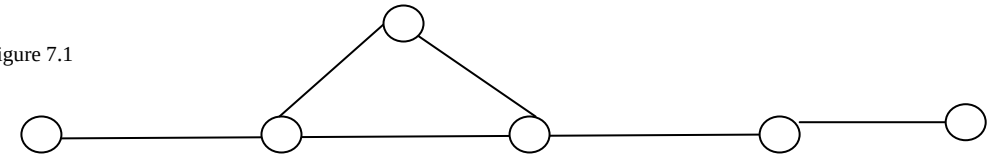
[71] As we saw in chapter 6, in the special case of spatial categorical properties, their identity is only 'partly' primitive. But there is a quidditistic aspect to them nonetheless.

regress, or 'infinite complexity' arising here, as every natural property has its identity fixed by another natural property, which has *its* identity fixed by a third property, and so on. But it need not be a regress -- the pattern of relations may turn back on itself, creating a network. Whereas with Armstrong's argument concerning the ontology of dispositions formed into a circle - a circle that looked to be vicious (which can be countered simply by asserting that dispositions are things in themselves), Bird claims, using a 'graph theory', that the kind of circularity involved in the fixing of properties' identities is virtuous. The thesis that needs defending is this:

> (S) The identity and distinctness of the elements of a set *E* of entities supervene on the instantiations of some relation *R* (or set of relations $<R_i>$) on the elements of *E*.(Bird 2007: 139)

When we translate this issue into graph theory (S) becomes '(S*) The identity and distinctness of the vertices of a graph can supervene on the structure of that graph'(*ibid*). What are required are graph structures whereby no rotation of the graph will swap the vertices and yet leave the structure of the graph the same. Any graph where this were possible would fail to fix the identity of the vertices. However, there are many graph-structures than do satisfy the criteria for fixing identity. Bird provides the following example (the circles denoting the vertices (properties), the lines denoting relations (second-order relations between properties) :

Figure 7.1



Aside from the graph with just a single vertex, the graph in figure 7.1 is the simplest non-trivial asymmetric structure that can fix the identity of its vertices. It follows that there must either be one, or more than five fundamental properties; so it would be impossible for there to be two, three or four fundamental properties. Bird does not see this as hugely problematic. He claims it is more than plausible that there is only one kind of fundamental property (where different non-fundamental properties

correspond to different structures of instantiation of the one fundamental property), and it may well be the case that, if there is more than one kind of property, they relate to one another in such a way that there must be more than five.

Within the context of graph theory it is possible to include loops and digraphs (directed graphs) to add further asymmetries. The *directness* of a relation represents the relationship between a property and its manifestation. Clearly there needs to be some directedness when using graph theory as a representation of the relations between properties and their manifestations are, in most cases, irreversible. Once we allow directedness into the graph theory the graphs look to provide an acceptable representation of the property networks, and the identity regress problem seems to be have been dealt with (as the identity of the vertices is uniquely fixed by the structure of the graph).

I have looked, then, at the three classic versions of the regress objection, and shown the responses to these objections as put forward by Bird and Bostock. As far as I'm concerned these responses are, at least on the face of it, sufficient to deal with the problems as raised by Armstrong *et al*. However, in the following section I return to the plausibility of the dispositionalist's relational constitution of identity, concluding that their ontology requires primitive identities after all.

7.6.1 Implicit Categoricalism in Neo-Dispositionalism

The identity regress is resolved by the relational constitution of dispositional properties – we can fix the identity of a property in the same way as the identity of a node in a graph can be fixed by its relations to other nodes in graph theory. But the essence of neo-dispositionalism is that the fundamental natures of all properties are wholly powerful; that is, (a) there are no categorical properties, and (b) the dispositional properties themselves are entirely devoid of quiddities. The SM-relations are second-order properties; that is, they are properties of properties. They play not only a role in fixing the identity of properties, but they also play an

important role in constituting the 'nature' of a property. It seems to me, then, that for the neo-dispositionalist, these second-order properties must also be powerful – they must therefore lack quiddities (primitive identity-fixers), having their identities fixed by some other means. Barker (2009) shows us that it is not clear, however, that the identity of these relations *can* be fixed in any other way than by invoking quiddities. If Barker is right, the very foundations of neo-dispositionalism are undermined – the dispositionalist requires all properties to have entirely dispositional natures, but as we will see, if they do, then the position is susceptible to an identity regress.

Barker provides two main arguments, the first of which I outline below. The second concerns the regress one encounters when we try to fix the identity of SM-relations. I present this in 7.6.2.

Dispositional properties in the 'powers' sense, are usually conceived of as properties of objects. An object, *x*, can instantiate the dispositional property charge -1, which, if possessed, amongst other things disposes it to accelerate away from positively charged object, *y*, when placed in the electrostatic field emitted by *y*. The dispositional property charge -1 is powerful because when instantiated by *x*, *x* has certain powers that it did not have previously; that is, it responds to certain stimuli it wouldn't have done if it did not have this power, and it can 'act' on objects instantiating properties of which charge -1 is itself a stimulus property. A dispositional property is powerful because it gives powers to the objects that instantiate it.

As we saw, according to the view outlined in the previous section, the identity of dispositions is fixed relationally; that is, fixed by the network of SM-relations. The SM-relations a dispositional property stands in to other dispositional properties are second-order properties, or properties of *properties* - dispositional properties (powers), on the other hand, are supposed to give powers to the particulars that instantiate them – they are properties of *particulars*: the first-order, natural properties. Thus

the relations constituting the identity of dispositional properties cannot be powers – at least not in the way powers have previously been conceived.

7.6.2 The SM-Relation Identity Regress

The second attack mounted by Barker begins by suggesting what would, if SM-relations were powers without quiddities (for Barker, a quiddity being a primitive identity-fixer), fix their identities. Unless powers are relationally constituted then we get an identity regress even at the first-order property level. However, not only natural properties, but also second-order properties must have their identities fixed by their place in an asymmetric graph. SM-relations, if they are to have their identity as powerful relations between natural properties fixed, must have their identities fixed by their relations to other properties. 'If [SM[72]] is a power, it is relationally constituted by SM-relations to other properties, which in turn must be relationally constituted by SM-relations, and so on' (Barker 2009: 3). We thus get a regress, and this regress is vicious. Furthermore, it's not clear to me what properties these second, third, fourth... order SM-relations could be, other than second, third and fourth order properties respectively[73].

7.6.3 Responses to the SM-Relation Identity Regress

Stephen Mumford suggested to me two responses to this problem: the first being that it is wrong for dispositions to be characterised in terms of SM-relations[74]; the second being the denial that relations are powerful - the neo-dispositionalist can, he claimed, avoid the problem by committing himself only to the claim that all *properties* are powerful, and deny that relations need to be powerful as well.

I agree with Mumford that the simple SM-relation characterisation seems inadequate, but this will not solve his problem. Mumford writes:

---

[72] Barker calls the relations SR-relations (stimulus-response relations) rather than SM-relations. But to be consistent I will continue to use SM-relation, as they are, as far as I can see, identical.
[73] See chapter 9 for more detailed discussion.
[74] This suggestion is presented in a forthcoming paper 'Mutual Manifestation and Martin's Two Triangles' (Mumford), to be published in a Jonathan Jacobs edited volume.

> But the process of production is depicted as an equal partnership, hence the view jettisons the Aristotelian idea that one partner is passive and the other active (a view that the stimulus-response model retains)... This may seem to be only a modest amendment to the stimulus-response model but we will see that it has quite radical consequences. But the first is simply that if the response (the effect, the manifestation) is produced by the partners working together, and neither could have acted alone, then we have a more equal and symmetrical relationship. It would be misleading, therefore, to label one partner the disposition and another partner the stimulus. In a way, each power is the other's stimulus, so which is the disposition and which is the stimulus is entirely relative to their perspectives. We therefore should dispense with stimulus talk as misleading. (Mumford: forthcoming)

I have no issue with these claims. In fact, I think to distinguish, ontologically, the 'stimulus' property from the 'disposition to be stimulated' property would be an obvious error - perhaps one that people have been making, but in my opinion not obviously so. My willingness to accept Mumford's point here is perhaps indicative of my unwillingness to accept it as a successful response to the problems posed in the preceding section. The mere fact that the disposition and the stimulus are reciprocal disposition partners does not dispose of the need for the higher-order SM-relations.

Mumford uses the example of ice coming into contact with water, and how this cools the water. I'm happy to accept that, ontologically, we should not distinguish the water from the ice in terms of which plays the active and which plays the passive role, but nonetheless, Mumford needs a second order property; that is, a relation between properties involved in the event – a relation linking the dispositions of the water and the dispositions of the ice. The manifestations may occur simultaneously (be mutual), but ultimately, under the dispositional essentialist analysis of causation, we still have a causal sequence that occurs in virtue of the nature of natural properties. There still needs to be a second-order relation between these properties – a relation that is part of the nature of each property linking them together; and it is the nature of this relation,

whatever it may be, that should be troubling the neo-dispositionalists. Now let us look at his second response:

Mumford suggests that the neo-dispositionalist can deny that relations are powerful. Barker takes care to prove that SM-relations are not powers, but apparently Mumford is happy to accept this, anyway (or at the very least, he deems it consistent with neo-dispositionalism). I think this raises the question, however, of how natural properties are to be free of quiddities, when the neo-dispositionalist is committed to their relational constitution?

The fact that first-order properties are constituted by relations to other properties (second-order properties) does not commit properties to being relations themselves (the nodes in the graphs are nodes, not relations), but nonetheless, these relations are still properties of the natural properties – they determine its nature.

What this means is that although the neo-dispositionalist's natural properties still have their *identities* tied to their causal/nomoligical role (and are hence distinctly non-categorical), the nature of each 'dispositional' property is inherently linked with quiddities. If this is the case, then neo-dispositionalism, and in fact dispositionalism in general is on the ropes. In the final chapter I will conclude that this forms the foundation of a knock-out blow for the neo-dispositionalist, forcing the powers theorists into admitting that theirs is a categoricalist metaphysic after all. Just as Lewis points out that being called Armstrong does not give you massive biceps, so it is true that calling a property 'powerful' does not make it so.

7.7 Dispositionalism and Laws

I now put these significant worries on the back-burner for the time being, with a view to returning to them in the concluding chapter.

Alexander Bird and Stephen Mumford *qua* neo-dispositionalists take laws to be aspects of reality; that is, laws are 'in the world', in their

metaphysical analysis. Mumford writes that ‘Laws must add something to nature such that the world would be significantly different were they not there’ (Mumford 2004: 145) – he claims that although the Humeans claim to have an account of laws, their ontology is in fact devoid of laws (as there is nothing in the world over and above the Humean mosaic). This, I take it, must stem purely from intuitions shared by Armstrong, but not of course by Humeans or Scientific Realists like Ellis. Similarly, Bird claims that ‘people… are willing to regard a proposition as asserting a law if...’ (Bird 2007: 203) – Bird thinks that people believe certain propositions *assert* laws, and not that certain propositions *are* laws.

Even though both Mumford and Bird agree that laws must be a part of the physical world, they disagree on whether there are any laws at all. Mumford holds a position that he terms ‘Realist Lawlessness’ (Mumford 2004), which is supposed to be an anti-Humean (in so far as it allows for metaphysical glue) position, but is also lawless in the sense that there are no real laws *in* nature.

Bird, on the other hand, sees laws as both existing, and also governing (although not extrinsically governing, in the Armstrongian sense). Given the discussion above, Mumford’s view seems to fall most naturally out of dispositional essentialist ontologies, but Bird’s view, I think, is plausible if dispositional essentialism is. In this section I will discuss both realist lawlessness and Bird’s view of laws in the neo-dispositionalist framework, and present arguments for and against each position. I will consider Mumford’s ‘Central Dilemma’, which purports to show that there are no laws in nature, and then demonstrate Bird’s response, which allows laws to be *in* nature, whereby laws supervene upon powers. Ultimately I argue that some of Mumford’s assumptions are unjustified, and that Bird’s account of laws as supervening on fundamental properties is far more appealing.

7.7.1 Mumford on Laws *in* Nature

Mumford's 2004 book is entitled: 'Laws *in* Nature'[75]. One might think that Mumford wishes to distinguish 'laws *in* nature' from 'laws *of* nature' (the usually terminology), in which case he might argue that there are laws of nature but not laws *in* nature. In fact, some Humeans might wish to endorse this – for some Humeans, laws are propositions[76], and although which propositions are laws supervenes on the pattern of instantiation of local particular matters of fact, they are not seen to be *in* nature *per se*[77]. However, Mumford seems to hold the view that any metaphysical position without laws *in* nature is a metaphysical view without laws, claiming that despite his dispositionalist metaphysic being a realist metaphysic (in so far as there is more to the evolution of the physical system than just one event after another), it is one completely devoid of laws - powers do all the work.

Mumford provides a number of arguments against there being laws in nature. I summarise the most interesting ones, below:

1. There is not, nor can there be, a nominal essence of laws (in nature)
2. There are disagreements as to what laws are *supposed to be*
3. 'Science does not give to its laws the same fundamental importance that is given to them by the metaphysical conception employed by nomological realism' (Mumford 2004: 134)
4. The Central Dilemma

[75] My emphasis.

[76] For Lewis, propositions are sets of possible worlds (See Lewis *On the Plurality of Worlds* Blackwell Publishing 1986 pp 27-50), but further discussion of this is unnecessary in the context of this thesis

[77] David Lewis might not endorse this view, as for him propositions are much like properties: 'the proposition is the same thing as the property of being a world where that proposition holds; and that is the same thing as the set of worlds where the proposition holds. A proposition holds at just those worlds that are members of it'(Lewis 1986: 53-54). Laws are, in a sense, in nature, as to have a proposition, *p*, as a law, is just to be a world that is a member of the set of worlds where (a) *p* is true, and (b) *p* is a law.

With respect to (1), it is true that a complete set of the nominal essences (that is, the properties we look for in a conceptual analysis) of laws is difficult to identify. In chapter 2, I considered various possibilities, including universality, mind-independence, and our ability to appeal to them in explanations and predictions. But there are certain fundamental, crucial aspects of the concept of law that people disagree on. Perhaps the most important of these is concerned with what laws actually *do*:

Do laws describe, or do they prescribe?

Do they govern and restrict the behaviour of otherwise inactive objects, or is it the other way around – do the laws supervene on movements of objects?

Are laws necessarily (*de dicto*) extrinsic to objects, or can they be intrinsic to them?

Could the laws have been otherwise?

With respect to (2), Mumford highlights the differences between the various metaphysical positions (Mumford 2004: 132), and in particular the differences between Ellis's and Armstrong's ontologies (but Bird's neo-dispositionalist account of laws could be added to this list). Armstrong sees laws as relations between universals (in nature), whereas Ellis sees laws as (propositions) describing the essential properties of natural kinds. Clearly very different conceptions of what a law is supposed to be.

In response to Mumford here I would simply say that both the Armstrongian view and scientific essentialism are attempts at metaphysical analyses of our concept of laws of nature. No doubt each is motivated by the variations in their own beliefs about what laws should be – Armstrong believes that laws must ultimately govern otherwise inert particulars, and Ellis believes that laws are closely linked with the highly structured world that we live in. Both, though, are trying to satisfy the rest of the conditions for being a law of nature, whilst maintaining a tenable

metaphysical position. Disagreements in the metaphysical nature of what it is to be a law are inevitable given the disagreements regarding what they see to be the nominal essences of laws, but nonetheless, I believe that if the conditions I identified in chapter 2 can be satisfied, whatever satisfies them is worthy of the name: 'laws'.

I don't see objection (3) as overly problematic, as it seems to me that arguably one of the main aims of science, and physics in particular, has been the discovery of laws and their subsequent applications (to space exploration, for predictions, and even in manufacturing). Perhaps Mumford's objection would be better put as his belief that attributing the *name* or *status* of 'law' to a proposition/equation seems more or less arbitrary in the realm of science. In science we find 'laws', 'theorems', 'principles', 'rules' and 'hypotheses', and whether a discovery is termed a law or a principle just doesn't seem to follow from any of the intrinsic properties of the propositions/equations in question. Bird, however, disagrees, insisting that there are real differences between laws and principles etc. In short, Bird believes that whether a physical principle is called a law or not depends on whether it fulfils certain criteria at the time of its discovery – the criteria essentially being the conditions arising from a conceptual analysis of laws. For Bird: 'Relative to a particular field and the state of knowledge concerning it, S is held to state a law if and only if':

1. It states a reasonably general relationship between quantities and properties;
2. It seems to state a relationship that is close to fundamental;
3. It is a new discovery, not easily deducible from known laws;
4. It has wide application in predicting and explaining phenomena;
5. The relationship seems necessary rather that accidental. (Bird 2007: 199)

When we look at lists of laws, principles and theorems, we see that, at the time of discovery, those statements that satisfy these conditions tend to have been called laws, and those that did not were not called laws.

Importantly though, it seems to me that whether or not scientists use the term 'law' consistently enough for the philosophers' liking should not overly concern us. No doubt all our beliefs about laws have in some sense been shaped by what the scientists call laws, but philosophers have the right to take what we can from science, and use it in metaphysical theory. Perhaps telling a scientist that 'such and such principle' is not a principle but a law would not be of concern to him, as the scientist is not metaphysically loading his discovery. However, we, as metaphysicians, can take the concept of law that has evolved from the scientists' work, and attempt to unravel the metaphysical underpinning of these laws. If we can find something that satisfies all the conditions we postulated in our conceptual analysis, then this is, as a matter of *de dicto* necessity, a law of nature - in the end, I do not think it really matters whether there's much consistency in the application of the term amongst scientists.

Mumford's main objection to there being laws in nature is (4): the Central Dilemma.

The Central Dilemma is presented as follows:

I. Either laws have [1] some (governing or determining) role or [2] not.

II. If [1] there is a (governing or determining) role for laws, then such laws are either [A] external to the things for which they play that role (they govern or determine) or [B] they are internal

III. If [A]: a theory of external laws is in need of an account of how laws relate suitably to the things they govern. The most plausible such account, the DTA theory [A_r], still has difficulty in explaining this relation and entails an incredible thesis (quidditism). Any other theory of external laws would need to avoid the problem of [A_r], but in a way that provides an account of A that is at least as plausible as [A_r].

IV. If [B]: A theory of internal laws is in need of an account of how laws could be suitably internal. The most plausible such account, reduction [B_r] (Ellis), is an implausible account of governance. Any other internal account would need to avoid the problem of [B_r] but in a way that provides an account of B that is at least as plausible as [B_r]. (Mumford 2004: 158)

He concludes that neither [A] nor [B] look good, and so we should reject (1); that is, we should reject that there can be laws that play a governing or determining role. This, combined with Mumford's realist metaphysics of dispositional properties, leads to the view he calls Realist Lawlessness.

7.7.2 Objections to Mumford's Arguments

The first objection I wish to mount against Mumford is once again concerned with the conceptual analysis of laws he presupposes. Let us grant, for the sake of argument, causal realism in the Mumford/Bird/Ellis sense – from this I think it reasonable to suppose that (although this is not necessarily so) laws, if there are any such things, either are, or describe, the aspects of reality that provide the metaphysical glue that maintains the uniformity of nature.

For Mumford, laws must play either a 'governing', or a 'determining' role – in this respect they must be *in* nature, not merely descriptions of it. Compare this with Ellis's laws: Ellis claims that laws *describe* the essential properties of natural kinds – if kinds have essential properties, then (at least according to Ellis) there are laws, as there are descriptions of the essential properties of natural kinds.

Mumford calls Ellis's view a 'reductive account', presumably because he sees Ellis as reducing laws to some other kind of thing. Thus, if Ellis can show this other kind of thing to exist, then he can show that laws exist. However, Mumford thinks this rules laws out from playing the roles they have to, to satisfy the criteria of being a law. 'How could something govern, or play any determining role in, that to which it is reducible?' (Mumford 2004: 155).

Mumford is right, I think, in proposing that nothing can govern something to which it is reducible, but does that really serve as an adequate critique of scientific essentialism. To begin with, Ellis is not reducing laws to essential properties of natural kinds. Laws are descriptive propositions for Ellis, not aspects of reality (although they express *de re*

necessities): they cannot be reduced to something physical. Laws, for Ellis, *describe* the essential properties of natural kinds. These descriptions clearly do not govern what the essential properties of the natural kinds are, as Mumford proposes laws must. But then, Ellis wouldn't want them to.

What properties are essential to a natural kind is a mind-independent aspect of reality, not governed by anything. If Ellis is right, and the world is structured in the way he supposes, then there are properties instantiated by members of natural kinds as a matter of *de re* necessity. Ellis's laws are global in scope, explain the regularities in nature, help us make predictions, and so forth. Laws, in Ellis's ontology, describe the aspects of reality that constrain it, and ensure that nature remains uniform. The world thus evolves in accordance with the *nature* of the things that constitute it: For Ellis, then, the laws tell us, but do not determine, what the natures of these things are. It seems to me that if Ellis is right, his laws are (a) very useful, and (b) would satisfy the conditions of lawhood in common with most people's conception of laws. I do not think, then, that the second horn of Mumford's dilemma is too problematic. He claims firstly that 'a theory of internal laws is in need of an account of how the laws could be suitably internal' (Mumford 2004: 157). Ellis (if one is happy to embrace scientific essentialism), I think, succeeds in providing us with such an account; and secondly Mumford argues that 'the most plausible such account, reduction (Ellis), is an implausible account of governance' (*ibid*) – I say "So what? It's by no means clear that we need an account of laws that govern, in order to have an account of laws!"

The second objection to Mumford's argument comes from Bird (2007). Bird claims that laws *could*, in fact, be seen to play a governing role. To reconcile a realist view of laws with neo-dispositionalism, Bird suggests a supervenience theory of laws, where laws supervene upon potencies. In the next section I will discuss Bird's suggestion, and how I think this leads us to a potentially catastrophic objection to dispositional essentialism.

7.7.3 Bird's Laws of Nature

Bird, like Mumford, thinks that laws have to be *in* nature in order to do the work they have to do. Laws cannot be propositions (although they can be expressed by propositions), but must be aspects of reality capable of determining how the system evolves - but they must also be, or be closely related to, the potencies of things (which are of course internal properties of things).

So what are laws according to Bird's neo-dispositionalist account? They cannot be the properties themselves, for to say a property in and of itself is a law would completely undermine our conception of law. We can, Bird claims, nonetheless claim that laws supervene on dispositions – but they should be seen as relations between properties rather than the properties themselves. This is formalised as (regularities):

V*: ∀x(finks and antidotes are to D are absent → ((Dx & Sx)→Mx); or

V** ∀x(*ceteris paribus* ((Dx & Sx) → Mx) (Bird 2007: 60)

Bird includes the condition of there being no finks or antidotes in order to overcome objections to the arguments regarding the non-universality of laws (see Cartwright 1995). However, given that, as I have shown previously, we have good reason to think that fundamental properties have neither finks nor antidotes, it seems to me that the *ceteris paribus* clause is surplus to requirements in the case of fundamental laws (which are relations between fundamental properties). Fundamental laws, then, do not require the *ceteris paribus* clause (note, though, that the formalisation would still work for fundamental laws, it's just that the *ceteris paribus* clause is redundant).

Although Bird agrees with Mumford as to the most fundamental explananda of the uniformity of nature and the truthmakers of counterfactuals being potencies, he concludes that there are laws in nature; laws that explain, hold universally, are mind-independent and non-accidental - laws are relations among properties that supervene on

potencies (Bird 2007: ch9). So again, Bird's neo-dispositionalist metaphysic of laws is closely tied with his metaphysic of causation - laws can be *expressed* as regularities, but the truthmakers of his law-statements are the same properties that provide the metaphysical glue in causal interactions. If Bird's metaphysics of causation is tenable, then on the face of it so is his account of laws, as it seems to satisfy the conditions to be found in all plausible accounts of laws of nature.

I have now presented the four main accounts of cause and law: the Humean account, the Armstrongian view, Scientific Essentialism and Neo-Dispositionalism. In the next chapter I will evaluate all four in the context of what I consider, from a metaphysical perspective, to be the most important principle in physical theory: The Principle of Least Action (PLA). It is so important because it has held true in all widely accepted physical theories since Newton. If any scientific hypothesis could be confirmed by observation, it would be the PLA.

Chapter 8: Consistency with our Best Scientific Theories

8.1 The Principle of Least Action

I claimed at the beginning of this thesis that the best understanding of our physical system will come from the advancement of both science and metaphysics. I will not go so far as to say that any metaphysical theory inconsistent with current scientific theory should immediately be discarded, as there are numerous examples of scientific theories that have been falsified - so why think our current scientific theories won't be falsified, too? (Newtonian forces, for example, look like Armstrongian governing laws – but Newtonian mechanics has been superseded by relativity theories). However, there is one scientific principle that has held true of all recent workable systems, including Newtonian Physics, Special and General Relativity, and String Theory. I refer to The Principle of Least Action (PLA), which states that for any closed physical system the action of that system will be minimised[78]. The form of the equations that determine the action of each system differ from one system to the next (these are called the Lagrange equations), but the basic notion of minimising action has always held true. Ultimately there is little doubt that ours is a world that evolves in accordance with least action principles, whatever the (using Ellis's terminology) fundamental laws of nature turn out to be.

If this is a justified claim (which I will just assume), it seems to me that any feasible metaphysics of laws must be able to accommodate it. In the following chapter I will firstly explain the PLA in more detail, and then look in turn at each of the metaphysical theories we have discussed (except the Armstrongian view which has already been discarded) to determine whether (a) they are compatible with the phenomenon, and (b) if the PLA is a law, how should it interpreted metaphysically, as opposed to purely a mathematical formalism?

[78] Or more accurately, 'extremised'.

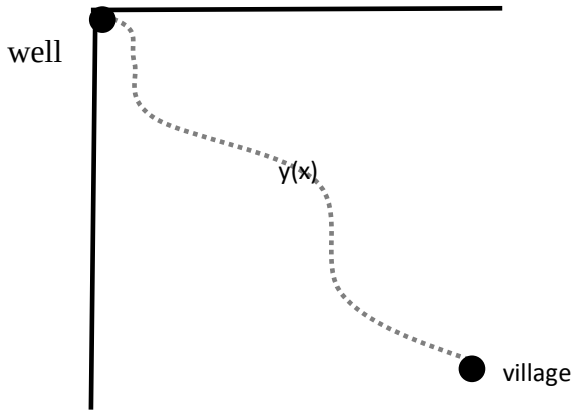
A detailed understanding of the PLA is not required for my purposes, so I think an analogy in terms of 'effort' rather than action would be sufficient, and less technical.

8.2 The Lazy Philosopher and The Principle of Least Effort

Let me introduce Matthew: an ex-philosophy student/turned lazy gardener. Because of the economic downturn and the British government's lack of appreciation for the value in the Arts, Matthew couldn't find an academic job straight away, and he was a little short of money - so after a brief failed attempt at joining the professional snooker circuit, he took to pumping water from a well into a Tibetan village to earn his right to live as a Buddhist monk. The well is already built, so all he has to do is lay the piping down to the water tank in the village. Of course due to his innate laziness, Matthew wants to spend as little time and use as little energy as possible to fill the tank, but as he soon finds out, the amount of time and energy required depends upon how he lays the piping. Initially he just laid the piping randomly from the top of the hills to the bottom, but it took a lot of effort to fill the tank, so he re-laid the piping and found the work a lot easier. Remembering his philosophy of physics and mathematical training, rather than trying every single possible route (of which there are an infinite number!), he realised the simplest way to choose would be to formalise the topology of the hills, and use a bit of calculus to work out the best route - by assigning a value to each point in the valley indicating the amount of effort it takes for the water to pass through the pipe when laid across it, he can find the 'effort function' which tells us how much work it takes to fill the tank for any route[79].

[79] This work is part of a joint paper with Karim Thebault. Thebault explains that 'the path of least effort will be the one for which very small variations in the path produce no change in E (the effort required). This is because finding such an extremal path indicates that the quantity E is stationary (either maximum, minimum or an inflection) and given the physical structure of the problem only the minimum option is possible. So the problem of our lazy plumber is solved by insisting that for variations between paths infinitesimally close to his test path there is no change in the amount of effort ($\delta E=0$)' (Smart, B and Thebault, K '*A Powerful Account of a Lazy World*' Presented at the 2010 British Society of Philosophy of Science Conference.)

The Plumber's Problem: figure 8.1



By formalising the topology of the hills and using the effort function, Matthew was able to find out how to fill the tank in the village in the quickest possible time (allowing him to meditate further on the time it takes to get a response from the top philosophy journals)[80].

Replacing *effort* with *action* (a notion closely linked with effort given its *SI* units are joules-seconds), our world, it turns out, is as lazy as Matthew, as the action is always minimized.

8.3 The Principle of Least Action

For a given physical system there is a special function particular to it called the *Lagrangian*, *L*, (When Mathew gave numbers to each point in the valley, the equation he used – but did not define – was the equivalent of the Lagrangian) which associates each possible worldly state of affairs with a number – and each possible state of affairs is represented by a point on the 'velocity-configuration space' graph. The Lagrangian 'is a function of all the… intrinsic properties ascribed to the objects in the system by

[80] Fortunately for Matthew, the institute of Applied Philosophy in Tibet heard of his endeavours, and immediately offered him a permanent position.

classical mechanics[81]' (Katzav 2004: 208) and is typically equivalent to the kinetic energy of the system minus the potential energy[82].

The action, *I*, like the effort in the case of the lazy philosopher, can once again be worked out using calculus once the Lagrangian and the path is known[83].

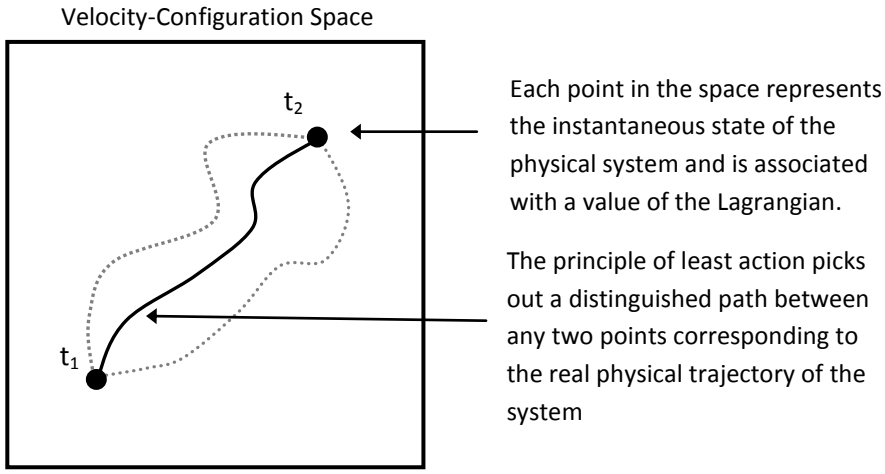


Figure 8.2[84]

As I have already stated, the PLA is not restricted to Newtonian Mechanics – it is found in Special and General Relativity, and even String Theory (the difference just lies in the form of the Lagrangians (in classical mechanics it is just kinetic minus potential energy, but in contemporary physical theories the Lagrange equations are more complex. Nonetheless, the principle still holds). The PLA seems to satisfy all our criteria for true

[81] i.e. space-time co-ordinates, velocities, charges and masses

[82] The Lagrangian relevant to general relativity represents an important deviation from this

[83] As Thebault writes: 'The *action*, *I*, of a physical system is then defined between two points in velocity-configuration space for any given path, γ, between those two points. A direct physical interpretation of action is not generally given in physical theory, but its SI units of joules-seconds indicate a close connection with both energy and time (and justifies our analogy with the plumbers concept of effort). The action can be explicitly calculated by the integral of the Lagrangian with respect to time along a path.
$I=\int_{\gamma} Ldt$' (Smart and Thebault: *A Powerful Account of a Lazy World*)

[84] Figure 2 shows the velocity-configuration space of a four-dimensional world, where each point represents the state of the world (including velocities, mass, charge etc) at a particular moment. The dotted lines represent possible paths that could have been taken to get from the first point to the second (sets of possible states the world could have taken en route to the second point). The path of least action – which will always be the actual path – is shown in black.

law-statements, and given its universality across all physical theories it is extremely important, I think, that we consider its compatibility with our metaphysical views[85]. Joel Katzav argues that the PLA is incompatible with dispositional essentialism – if he is right, I agree with him that dispositional essentialism is in serious trouble[86].

There are two objections Katzav raises, based on his assumptions about what the PLA presupposes. I will outline the two presuppositions (8.4.1 and 8.4.2), and then see how each of the metaphysical theories can tackle the supposed problems[87].

8.4.1 Many Possible Paths

Katzav believes the PLA allows that the history of a physical system *could* have corresponded to a velocity-configuration space path *without* a minimised action (there were many routes Matthew could have chosen for his pipeline), and from this he makes the non-trivial assumption that PLA presupposes its own contingency. He reasons, then, that any theory entailing the possibility of only one set of laws *must* be incompatible with it.

8.4.2 The Explanatory Power of the Principle of Least Action

Katzav argues that the Principle of Least Action is the ultimate explanation for why closed physical systems evolve in the way that they do.

The PLA tells us that the physical system's quantity of action must be the smallest relative to the quantities of action it might have possessed;

[85] Arguably the PLA should not be thought of as a law itself, but as a principle that describes the way our physical system evolves. In these terms an adequate account of laws should explain the PLA but need not include it as one of its laws.

[86] My thanks to Karim Thebault, with whom I have written a paper entitled *A Powerful Account of a Lazy* World (recently presented at the 2010 British Society of Philosophy of Science conference). The contents of this section form a part of that paper.

[87] Note that Katzav presents these as problems for Ellis's dispositional essentialism (this is distinct from neo-dispositionalism and of course from Humeanism).

as stated earlier, if we know the system's action integral, the actual equations of motion can be calculated from this information alone. So the property of being a physical system that has a quantity of action that is a minimum can be appealed to in explaining why certain equations of motion describe the objects that comprise it.

Katzav then argues that, as the equations of motion describe the dispositions of the objects within it, and that physical systems having an action that is minimized (that is, that the physical system complies with the PLA) explains why those systems have the equations of motion they do, we should realise that the PLA also explains why the objects in the system have the dispositions they do. One might ask why these deductions are explanatory? It is, he argues, because 'that some quantity is [a minimum] seems to imply that, if it is actual, it is not an accident' (Katzav 2004: 215), and that something is not an accident allows us to appeal to it in explanations.

8.5 Can the Three Metaphysical Theories Accommodate Many Possible Paths?

I will now start the process of assessing each of the three main metaphysical theses against the problems posed by these problems, starting with the sophisticated regularity theory.

8.5.1 The Sophisticated Regularity Theory and the MPP

When we ask whether the system could have followed a different path, we are simply asking whether it would have been possible for our physical system to have evolved in a different way; that is, might certain events have occurred such that the principle of least action was violated? For the Humean this is certainly possible. The laws of our world supervene upon the spatiotemporal location of local particular matters of fact, and not *vice versa*. The regularities are not governed or determined by the laws, rather the laws fall out of the regularities and the best-systems account. In the

actual world the principle of least action is (we are assuming) a law *in virtue* of its being true, but it need not have been a law at all. The Sophisticated Regularity Theorist, therefore, has absolutely no problem in accommodating the apparent presupposition that the PLA holds contingently – it is a consequence of their metaphysical position that it must hold contingently.

8.5.2 Scientific Essentialism and the MPP

If Katzav is right in thinking the PLA presupposes its own contingency, then Ellis is in trouble. According to scientific essentialism and the dispositional properties fundamental to the position (assuming determinism) there is only one way in which our physical system could evolve given the initial conditions. The properties instantiated by objects have their dispositional natures essentially, and given the absence of finks and antidotes they cannot fail to manifest once stimulus conditions are met. Katzav argues that as this theory entails the possibility of only one set of laws (and thus only one possible path between the states of affairs t_1 and t_2 in figure 2), dispositional essentialism must be incompatible with the PLA.

For the scientific essentialist, it is true that the PLA is a law applying to all physical systems. It is not logically necessary that the PLA holds in all worlds, but it is *de re* necessary in our world in the same sense that electrons are necessarily negatively charged. According to Ellis, physical systems are Lagrangian (adhere to the PLA) as a matter of *a posteriori* metaphysical necessity: 'like accidental generalizations, [the PLA is] *a posteriori*, and can be established only by empirical enquiry, but unlike such generalizations, they are not contingent' (Ellis 2009: 65).

For Ellis, as with Katzav, it is no accident that the PLA holds in our world. However, unlike Katzav, Ellis takes it to be an essential property of the actual world, so its holding is not a contingent matter. From the quotation above we can see that this is true of all physical systems, for that matter, and although it is logically possible for other

worlds not to adhere to the PLA, according to Ellis none of those worlds would be physical systems.

Ellis sees the dispositional essentialism Katzav is *really* addressing to be what he terms a naïve form of the essentialist position (essentially the position I name (rather less patronizingly) 'neo-dispositionalism'), and in response to Katzav suggests that his more sophisticated version (SDE), which employs essential properties of natural kinds, can accommodate the PLA without difficulty.

He responds to the MPP by asserting the principle of least action to be 'of the essence of the global kind in the category of objects or substances' (Ellis 2005: 90), making it a law of nature that all natural kinds in that hierarchy adhere to the principle[88]. Any world in which the PLA does not obtain cannot be a member of the global kind our world *must* be a member of. It is thus *a posteriori* metaphysically necessary that there is only one metaphysically possible path, so it simply *cannot* presuppose its own metaphysical contingency[89]. Ellis thus gets around the objection by denying Katzav's presupposition that the PLA holds contingently.

8.5.3 Neo-Dispositionalism and the MPP

Bird and Mumford do not appeal to essential properties of natural kinds in their ontology, and so they are unable to give exactly the same response Ellis provides. Of course these neo-dispositionalists will face the same objection; that is, the need for the path of least action to be just one of a myriad of possible paths.

On the face of it Bird does seem to have a reasonable response. He claims the fundamental flaw in Katzav's reasoning is his implicit assumption that because something is logically possible it must also be metaphysically possible (see Bird 2007: 200-214). Although neo-

[88] Note that this does not commit him to the PLA being a powerful property of the objects
[89] It turns out there *are* a number of possible paths, but this possibility is only a logical possibility, not a metaphysical one; but more on this later.

dispositionalism rules out the *metaphysical* possibility of equations other than actual ones describing our system, it does not claim this to be a *logical* impossibility. But this is not an inconsistency; metaphysical impossibility does not entail logical impossibility, since according to dispositional essentialism, the domain of what is logically possible includes propositions which cannot be instantiated in any metaphysically possible world.

8.5.4 Conclusions Regarding the MPP

Each of the three metaphysical theories seems to have a response to the MPP problem consistent with its ontology: Any regularity theorist can claim that a law/regularity like the PLA holds contingently, and so they cannot be troubled by Katzav's claim; In Ellis's ontology (if acceptable) the PLA is a law, and hence it must be an essential property of the actual world - Katzav's belief that the principle presupposes its own contingency must, according to the scientific essentialist, just be false; and finally the neo-dispositionalist agrees that the path followed is *logically* contingent, but that this is not inconsistent with it being metaphysically necessary (which is all the dispositionalist is committed to).

Ultimately I don't think the many possible paths problem particularly undermines any of the theories. However, there are reasons for doubting the legitimacy of all three responses:

1. Anti-Humeans would argue that it is absurd to conclude that the PLA is a cosmic coincidence;

2. Those who, like myself, think we have no reason to believe that the laws of our world could not have been different, will question Ellis's response; and

3. for the neo-dispositionalist, although he explains why only one possible path is available, he does not explain why this path is the path of least action.

The objections to the Humean and Ellisian responses are really just reformulations of the objections to the positions themselves, but the objection to the neo-dispositionalist here is more interesting. Although Bird *et al* provide us with metaphysical necessity in the form of powerful properties, constituted by SM-relations, the principle of least action does not look to take this form. If neo-dispositionalism is correct, then all the fundamental dispositions must somehow 'add-up' to ensure the path of least action is followed – the path supervenes on the dispositions. But we have not been provided with any reason to think that this should be the case. Surely it can't be a cosmic coincidence, as eliminating cosmic coincidences is arguably the main aim of dispositional essentialism.

This leads us neatly into looking at the kind of explanations the PLA and our metaphysical theories afford.

8.6.1 Sophisticated Regularity Theory and the Explanatory Power of the PLA

The principle of least action is a global law of the highest order. For the sophisticated regularity theorist, this would translate as it being one of the fundamental axioms in the best systems analysis; from this principle and the other fundamental axioms (all expressed in terms of regularities) all the laws of nature can be derived, and, as I have shown (section 3.3.5), these regularities can act as explanations. Of course for the regularity theorist the PLA is not itself explained, but this is not problematic - explanatory chains always have to end somewhere, and this fundamental regularity is a plausible option for the ultimate unexplained explainer.

The regularity theorist would see the PLA as the ultimate explanans, as it is arguably the highest-order regularity when it comes to the evolution of our physical system. But this is not why Katzav thinks we can appeal to it in explanations. From the quotation above, we see that he thinks that as the path followed by the physical system is *precisely* that which minimises action, we should infer that it is no accident that it does so. In the context of Katzav's paper, 'no accident' cannot be a Humean

interpretation of ‘no accident’: he is implying that there is some metaphysical glue ensuring that this path is followed, as mere physical necessity would make the claim trivial (if we assume the PLA to be a law). This of course means that Katzav would never be won over by the Humean, even though Humeans can easily provide a response in terms of the best system analysis.

In short, the PLA would certainly explain the lower order regularities for the sophisticated regularity theorist. If the claim is that to be consistent with the PLA, the laws of motion, which are ultimately lower order regularities derivable from the PLA, must be explained by the PLA (and not vice versa), then, on his terms, the Humean has a perfectly adequate response, whether or not Katzav will acknowledge this.

8.6.2 Scientific Essentialism and the Explanatory Power of the PLA

As Katzav seems to present it, the argument against Ellis’s version of dispositional essentialism runs as follows:

> P1. The quantity of action in the physical system being an extremum explains why certain equations of motion describe the way the objects that comprise it behave.
>
> P2. The equations of motion of a physical system describe the dispositions of the objects within it.
>
> C. The quantity of action in the physical system being an extremum explains why the objects within the system have the dispositions they do.

Assuming that arguments of the form: *a* explains *b*, *b* describes *c*, therefore *a* explains *c* is valid, why is the PLA incompatible with dispositionalism in virtue of C? It is incompatible (according to Katzav) because for the dispositionalist, dispositional properties alone are the explanans for all the dispositions of objects (the way they behave in certain circumstances). If we accept that the PLA plays an explanatory

role, then there are explanations for physical phenomena independent of the dispositional properties of particulars. To put it another way; if we accept, (as we should, according to Katzav, on account of the fact that the PLA can be no accident) the vast explanatory value of the PLA, the dispositions of objects (the way they behave in certain circumstances) have an explanation that does not derive from the 'powerful' dispositional properties those objects instantiate, and this supposedly directly contradicts the essence of dispositionalism.

One might argue that not only can the equations of motion be derived from the PLA, but the PLA can be derived from the Lagrange formula and equations of motion of any physical system. Why not suppose *these* deductions are the explanations; that is, that the PLA is explained by the laws of motion?

Katzav anticipates this response, and responds that 'as the fact that physicists typically use the [PLA] to deduce corresponding equations of motion and not vice versa illustrates, the explanation only proceeds from the [PLA] to the equations of motion.' (Katzav 2004: 217). But, he claims, even if a deduction of the PLA *is* explanatory, we would still have reason to think that other principles (from which the relevant formulation of the PLA is deduced), are not compatible with dispositionalism.

I think Ellis can respond to the explanatory problem using the same principles as his response to the Many Possible Paths problem, in that he can focus on his formulation of scientific essentialism. For Ellis it is *de re* necessary that our world has the PLA as a property, as the PLA is an essential property of the global kind our world is a member of. According to Ellis, in science we should often consider 'X is an essential property of Y' to have genuine explanatory value. Just as we can explain all electrons having negative charge by appealing to 'negative charge' being an essential property of electrons, so we can explain our world evolving in accordance with the PLA in virtue of our world being a natural kind with the PLA (a natural kind of process) as an essential property. So

the PLA certainly does have explanatory value (the fact that it is *de re* necessary that our world evolve in accordance with the PLA seems like a reasonable explanation for the phenomenon!).

But that does not mean that other essential properties cannot have explanatory value, too (including dispositional properties). It is an essential property of electrons that they repel other things with negative charge, and we can appeal to this when explaining why electrons repel things with negative charge. In our world it might be entailed by the PLA that electrons repel things with negative charge, but the PLA is not an essential property of *all* worlds (there are many worlds where the PLA does not hold). Explanations in terms of essential properties of natural kinds of processes and substances, however, can hold across *all* worlds. If it is an essential property of electrons that they repel things with negative charge, all electrons in all possible worlds repel things with negative charge: 'it is an essential property of electrons that they repel negatively charged things' can, it seems to me, have explanatory value in certain contexts.

I claim, therefore, that the Scientific Essentialist can both explain the holding of the PLA in light of it being of the essence of the actual world, and (in agreement with Katzav) appeal to it when explaining the evolution of the physical system. Again, in agreement with Katzav, scientific essentialists can even appeal to the PLA as an explanation of why the actual world has the natural kinds of substances, processes and properties it does, but I claim he can also appeal to the essential properties of these natural kinds as explanantia in their own right. The supposed inability of the dispositions of things to explain the PLA is not significantly problematic for Ellis, as for Ellis there is more to explaining why our physical system evolves as it does than just the dispositional properties instantiated in the world.

8.6.3 Neo-Dispositionalism and the Explanatory Power of the PLA

So once again Ellis can fall back on his ontology of natural kinds, but the neo-dispositionalist cannot do this. For him, it seems to me, the entire evolution of our physical system really is entirely attributable to the dispositional properties of things (and nothing else). Although Bird can deal with the MPP problem, it is not entirely clear how he is to explain why the principle of least action holds in our world (even if he can explain why there is only one metaphysically possible path). Bird writes that 'the PLA is an *a posteriori* tool for providing the answer [to the question of which path will be taken]. That is consistent with the PLA itself being necessary, with the actual path being necessary, and with those necessities flowing from the (in this case unknown) intrinsic properties of the initial state of the system' (Bird 2007: 215). It is clear that through Bird's ontology the PLA falls out of the dispositional properties of our world and its initial state, but from what I have said about neo-dispositional so far we have no reason to think this is anything other than a cosmic coincidence under Bird's ontology; the necessities flow from the intrinsic properties of the initial state of the system, but presumably there are many ways the initial state of the system might have differed.

Bird deals directly with least action principles twice in *Nature's Metaphysics*; initially to dispel the many possible paths problem, but later when he refutes Mumford's claim that in science, 'relatively arbitrary and almost certainly non-metaphysical considerations…lead to something being called a law' (Mumford 2004: 134), Bird also explains why scientists label the PLA a *principle* and not a *law.* As we saw in section 7.6.1, Bird thinks scientists really do differentiate between laws, principles, and theorems, providing a list of conditions for laws. Bird argues that least action principles do not satisfy all these criteria: in particular they fail to satisfy conditions (2) and (3) – they do not 'state a relationship that is close to fundamental', and they are not 'new discoveries that are not easily deducible from known laws' (Bird 2007: 199). This, he claims, is because although the least action principles do

provide an alternative mathematically elegant formulation of the laws, we are not able to derive anything more than we were able to using what Katzav refers to as the 'laws of motion'. (Using Newtonian mechanics as an example, we have Newtonian laws of motion and the least action principle reformulation where the quantity of action is kinetic minus potential energy. The least action principle approach does not provide us with any more information, in the Lewisian sense, than do the Newtonian laws of motion).

Secondly, he claims that the kind of explanations the least action principles afford are 'explanatorily odd. In the case of Fermat's principle, how does the light ray know which path is the quickest?' (*ibid*).

In response to Bird's first objection, it seems to me that (2) *is* satisfied by the PLA. What could be more fundamental than a principle (or law) that has been true of *all* reasonable formulations of the evolution of our physical system?

Criterion (3) can be split into two parts: (i) that it is a new discovery, and (ii) that it is not easily derivable from known laws. With respect to (i) I do not see why being a new discovery is required for or relevant to a true statement being given law-status. Can what was once considered a true law-statement cease to be a true law-statement over time despite it never being falsified (even through numerous scientific revolutions)? And with respect to (ii), this presupposes what the known laws are. Following Katzav, why not suppose that the PLA is the known law and derive the 'laws of motion' from that and the relevant Lagrangian. Doing the work this way round would give the laws of motion 'non-law' status (which, I hasten to add, is not a conclusion I'd wish to make either).

Bird's second argument, that the least action principle explanations are explanatorily odd, is not exactly a knock down argument in the context of scientific theory. I agree that this kind of explanation is strange, but so are many explanations in science. Explaining light bending around stars by appealing to the curvature of spacetime is odd; the fact that time travels

slower for things travelling at higher velocities is odd; the M-Theory, (which asserts that the strings in string theory are really 1-dimensional slices of a 2-dimensional membrane vibrating in 11-dimensional space[90]) is odd; and the idea that a quantumly entangled particle can have a spin state that is neither up, nor down, nor neither, nor both is very odd... but these hypotheses serve as explanans for our empirical evidence, so the fact that the kind of explanation the PLA affords is explanatorily odd should not be too problematic. If I ask the question "why did our physical system follow this particular path through velocity-configuration space", it seems to me that "because that was the path of least action" would be a perfectly good explanation – and in particular, it would be better than "because of law of motion-1, law of motion-2, law of motion-3 and the initial state of the system". Perhaps this is due to an intuition that the Ramsey-Lewis 'simplicity and strength' criterion is a good way of judging between rival theories, but nonetheless this looks right.

8.6.4 Should the PLA be a law for Neo-Dispositionalists?

I try to show above that the criteria set by Bird for a statement to be a law-statement are either satisfied by the PLA, or should not be one of the criteria at all. But this does not mean the PLA will be a law for Bird, given his metaphysical conception of laws. Next I show why the PLA cannot be a law for the neo-dispositionalist.

As we saw in the preceding chapter, Stephen Mumford argued for a metaphysical position with metaphysical necessity grounded by dispositional properties, but devoid of any laws, as he assumes the governing role to be essential to 'real' laws (as we saw earlier (Beebee 2000), it is not at all clear that laws need any governing aspect to them at all). Bird claims, though, that the dispositionalist ontology is not incompatible with governing laws *tout court*, only *extrinsic* governing laws (in the Armstrongian sense). He writes that 'laws can govern even if they are potencies' (Bird 2007: 197), concluding that 'laws are general

[90] http://en.wikipedia.org/wiki/String_theory - I do not pretend to understand what this means, but more than 4-dimensions in spacetime seems odd to me!

relations among properties that supervene on potencies, and which have explanatory power' (Bird 2007: 200).

The PLA does not fall into this category, though, as although according to Bird its holding supervenes on the general relations among properties that supervene on potencies (in our world), it is not itself a general relation among properties with dispositional essences (potencies). So under Bird's ontology the PLA is not a law. I try to provide accounts what the PLA might really be, in the next section.

In conclusion to this section - Katzav challenges the dispositional essentialist to show how dispositional properties can play an explanatory role when it is clear (as far as Katzav's concerned) that this role is fully occupied by the PLA. As the PLA plays the role the dispositional properties are supposed to play, and the PLA is more fundamental, Katzav believes dispositional properties are made redundant. Bird refutes Katzav's claim by saying that the PLA is not a real law, but a principle that plays no metaphysical role in the evolution of the system. The PLA, for Bird, is just a neat mathematical formalism that captures and describes the motions and changes determined by the intrinsic properties of particulars. As it is just a principle and *not* a law, the PLA carries with it no genuine explanatory power (in the same way as, for Bird, Humean laws cannot explain because they are not metaphysically meaty), and so the challenge Katzav raises, that dispositional properties cannot play an explanatory role, is overcome. For Bird, all the laws of motion supervene on potencies.

This does not, however, solve the problem of why the PLA holds in the first place. We can see that for Bird the PLA holding in our world ultimately supervenes on potencies. But as Katzav says, its holding (at least for the necessitarian) makes it look like no accident that it does. Bird has still not explained why it is no accident that the *path of least action* rather than any other path is followed, and I'm sure that as a necessitarian he wouldn't want it to be accidental either. I will try to show in 8.7.4 that a

dispositional account of the PLA is not entirely out of the question, but whether it solves the neo-dispositionalist's problem is another matter.

8.7 So what *is* the PLA?

Katzav is trying to persuade us that the PLA explains the laws of motion and the dispositions of things (and not the other way around). If Katzav is to support his claim within a metaphysical debate he needs to give a *metaphysical* account of what the PLA actually *is*, which he fails to do in his 2004 paper. The PLA is, for the physicist, an alternative way of formalising, mathematically, the evolution of the system. This is not metaphysically informative. Katzav needs a metaphysical interpretation of this alternative formalism that differs from the metaphysics of the dispositional essentialist. There are, as far as I can see, several options as to what the PLA could be, each of which falls naturally into one of the three metaphysical theories this thesis is concerned with:

1. The PLA is just a true law-statement – expressible as a regularity.
2. The PLA is an extrinsic governing law that determines how otherwise inert objects behave.
3. The PLA is a dispositional property of the entire physical system.
4. The PLA is a dispositional property of all the constituent parts of the system.

I will deal with these options in turn:

8.7.1 The Regularity View

If we embrace Humeanism we could almost certainly interpret the PLA as one of the axioms in the Lewisian Best System. It would simply be a true law-statement expressing a universal regularity, and a very informative one at that - all the laws of motion would be derivable from it and the other fundamental axioms), and it would have explanatory value (at least as far as the regularity theorist is concerned). Taking the PLA to be just a

regularity, however, would take away the metaphysical necessity I think Katzav wants it to have.

We do not have to assume there is no metaphysical necessity even if we accept the PLA to be just a regularity. Bird, for example, might claim that the PLA is just a regularity, but it would be a regularity in virtue of the myriad of dispositional properties that constitute the world. This would, however, also strip the PLA of all its explanatory power as far as Katzav is concerned. Nevertheless, even if Bird is not happy to, the Humean in particular would be happy to embrace the PLA as a law.

8.7.2 The Governing PLA?

Option 2 has strong parallels with the Armstrongian view of laws of nature. According to this interpretation of the PLA, in every closed system there would be an extrinsic governing law determining what all the dispositions of that system must be (note that for Katzav, dispositions must just mean stimulus-response regularities, rather than some metaphysically stronger notion. I use disposition here purely because Katzav himself does). With Armstrong we saw that where N(F,G) holds, all Fs must, with metaphysical necessity (metaphysical glue) be Gs - Fs and Gs are both causally inert, and the N-relation is a universal that holds between these properties only contingently at any world. Where the PLA holds non-accidentally under (2), there must be some extrinsic property (that holds contingently at any world), P, ensuring that *all* the right regularities (or dispositions in Katzav's sense) hold omnitemporally, omnispatially, and with metaphysical necessity (involving connecting as a matter of 'contingent necessity' whatever sets of universals need to be connected). Property P (when it holds) determines what properties the N-relations stand between. The natural way to see this is as a third order relation between second order universals, rather like the third order N-relations used by Armstrong when explaining functional laws.

As the 'dispositions' of properties ultimately hold *in virtue* of the P-relation and the corresponding N-relations holding between the right

universals, the properties (as governed by the N-relation) are themselves inert, so just as with Armstrong, this externally governing conception is susceptible to the quidditism objections. But I'm not sure this alone would overly concern Katzav. After all, he thinks the PLA presupposes its own contingency, and so the idea that it will not hold in many possible worlds is beneficial (just as Armstrong with the N-relation, Katzav would think that when a world seems to be evolving in accordance with the PLA, inference to the best explanation will direct us towards the P-relation holding at our world – he says that if it holds, it holds non-accidentally). Katzav must, however, give some account of what the P-relation is. Under the governing law metaphysical conception, the P-relation would be a third order relation between universals – itself a universal; but as Bird (2005) showed us, this generates a vicious regress in the light of the quiddistic nature of universals in a metaphysics of extrinsic governing-laws.

8.7.3 A PLA Disposition Intrinsic to the System's Constituents

Only in the trivially simple case of a one particle system do objects individually follow space-time trajectories which are strictly speaking paths of least action[91]; the PLA applies to the action of the *entire system*, so generically all the trajectories of all the particles in four dimensional space-time must be coordinated such that the system as a whole follows a path of least action in velocity-configuration space. Why each *individual particle* follows a space-time trajectory corresponding to the minimum action for the system it is part of cannot, therefore, be explained purely in terms of a PLA disposition of any particular (unless in a single-particular system) considered in isolation.

The manifestation of the PLA disposition is not the object instantiating it following a trajectory that would, if all the other objects in the system followed suit, minimise the system's action, but the *actual* minimising of the system's action. This can be ensured by including the

[91] In fact, the trajectory for an individual particle in a composite system can be radically different to the trajectory corresponding to its least action if considered on its own.

other objects following 'minimising trajectories' amongst the PLA disposition's stimulus conditions. Furthermore, as the minimising of action is not restricted to any finite time period the manifestation of the PLA disposition must be the minimising of action over the entire duration[92] of the system's existence. Consequently the PLA disposition must (in a PLA world) be *constantly* manifesting for the PLA to be true in virtue of the PLA disposition. Furthermore, the manifestation of the PLA disposition instantiated by X requires all other objects to be following paths of least action, but if the other objects are following paths of least action independently of the PLA disposition, then the system's following the path of least action would (mostly) not be *due* to the PLA disposition (and we want the PLA disposition to be doing the work). We can solve this problem however, but ensuring that *all* objects are following paths of least action because of the PLA disposition they *all* instantiate.

For the stimulus conditions to be met, then, *all* objects in the system must take space-time trajectories that allow the system's action to be minimised. So given: i) the non-accidental nature of the PLA; ii) the fact that according to dispositional essentialism, the trajectories of objects are determined by their dispositional properties alone; and iii) the fact that the PLA disposition can exist only where the manifestation relata exists (since if all the objects in the system are to have the requisite trajectories then all of them must instantiate the PLA disposition), in every metaphysically possible world in which the path of least action is followed due to the PLA disposition, if the property is instantiated by one object, it is instantiated by all objects.

One might be tempted to object that if we accept the account provided above, overdetermination looks to creep into our causal picture - both the disposition of an electron to move towards a positively charged metal sheet *and* the PLA disposition are responsible for its motion. I don't see this as particularly problematic, as often more than one disposition determines the motion of an object. However, it does seem as if a cosmic

[92] Or over the relevant region of space-time in the case of a field theory

coincidence is required for this situation to arise – if this is to be a PLA world in virtue of the PLA disposition, every particular, omnispatially and omnitemporally, would need to instantiate the PLA property, and cosmic coincidences of this kind are not welcomed by the neo-dispositionalist.

Nonetheless, if these accounts are plausible, then Katzav's argument fails. The PLA does not *wholly* determine the laws of nature – they are partly determined by the other properties instantiated by particulars. But this should not be surprising. The PLA alone can never be sufficient to determine the evolution of the system, as we also need some means by which to determine what the quantities of action will be, and it seems to me that the other properties instantiated by particulars in the world are suited to playing this role – it is these properties that determine the Lagrange equations.

We can see the PLA in two perfectly legitimate but very different ways. The first, and this, I think, is how the scientists will generally see it, is as a way of mapping the evolution of the physical system – it is an alternative means (alternative to the more conventional laws of motion) by which to mathematically formalise the way closed systems behave. This is not metaphysically loaded.

The second way is to see the PLA as a law as (some) metaphysicians try to see laws, rather than just some mathematical equations. I provided several options: a truth that supervenes on local particular matters of fact, an extrinsic governing law, an essential property of the actual world, and a neo-dispositional account. To ask whether the metaphysical accounts central to this thesis are compatible with the PLA, then, is to ask whether the holding of least action principles (in the first sense) can be explained by each of the metaphysical theories.

For the Humean, the PLA would probably be the highest order regularity – the regularity which tops the explanatory chain as I described it in chapter 3. It cannot itself be explained in terms of its holding with metaphysical necessity, but of course the Humean wouldn't want this,

anyway. In a PLA world the PLA will be law – it would be a non-accidental regularity derivable from the axioms that comprise the best combination of simplicity and strength, purely in virtue of its *being* one of these axioms, and as such it has explanatory value, holds universally, holds contingently (thus saving it from the MPP problem) and can be used for future projections. Admittedly those unwilling to accept that the laws supervene on their instances rather than the other way around will be very unhappy with this view, but this is hardly the Humeans' problem - they do not share such intuitions. Least action principles, then, pose no problem for Humeanism.

The Armstrongian could accommodate the PLA in his ontology by taking the PLA to be a third-order property. Although compatible with the Armstrongian view, this interpretation of the PLA should be rejected on account of the vicious regress identified by Bird (2005).

For Ellis, the fact that our world evolves in accordance with least action principles can be attributed to the fact that if the least action principles did not hold, then not only would it not be our world, but the non-PLA world would not even be a physical system. It can, he argues, be established as an *a posteriori* necessity that the path of least action is followed by all physical systems (and so Katzav is just wrong in thinking the PLA presupposes its own contingency). Ellis's view is compelling in as much as the PLA does come out as a law of nature as it is an essential property of the world, but at the same time it does not undermine his claim that the motion of physical objects is entirely determined by the myriad of dispositional properties in his ontology.

The neo-dispositionalist has a little more work to do if the holding of the PLA is not to come out as coincidental. If the dispositional properties do not include any commitments to physical systems following paths of least action, then there seems to be no real reason why the dispositions could not have been such that this path was not followed, which of course begs the question of why it *is* followed in our world (for

surely they would not want to concede that we just *happen* to be in the PLA world). I have provided an account whereby we can accommodate the PLA by including a 'PLA disposition' with the stimulus condition being that all particulars in the system instantiate the property - this, I claim, solves the problem for the dispositionalist, but at the expense of a cosmic coincidence. We end up with a position that is not susceptible to Katzav's MPP problem on the grounds that the PLA presupposes only logical contingency, and not susceptible to the Explanatory problem as the PLA is itself a power amongst the many required to explain the evolution of our physical system. Nonethless, it is still not particularly appealing from a necessitarian's perspective, as avoiding cosmic coincidences is a major motivation for their view.

8.8 A New Approach

In the preceding sections I have been looking for a single metaphysical view that captures the least action principles, but it seems to me that this is the wrong way of going about it. The PLA and the traditional laws of motion are two formalisms of the same phenomenon – namely, the evolution of the physical system. These are mathematically equivalent, but the best metaphysical interpretations of the two might not be compatible. That is not to say that we must choose one over the other. If we want *de re* necessity, it seems to me that the following are the best options:

8.8.1 The Best Necessitarian Metaphysical Account of the Laws of Motion

We can rule out the Armstrongian view for the numerous reasons I have already outlined, and as the two remaining possibilities are the regularity theory and theories involving dispositional properties (and regularity theory is devoid of *de re* necessity), we are forced into a dispositionalist account (as a best necessitarian theory).

Introducing the PLA disposition seems a little *ad hoc* and unnecessary when considering the individual laws of motion – that is, the laws independent of one another. Introducing the PLA disposition was

something we had to do to try and explain the evolution of the entire physical system in accordance with the PLA. If we are just looking for a reason for the traditional laws of motion to hold, neo-dispositionalism without the PLA disposition seems to do the trick. It is in virtue of the dispositional properties in the world that the system evolves as it does - the dispositions have stimulus conditions and manifestations, and each causal interaction can be cashed out in terms of the manifestation of one or more dispositions. Similarly, individual causal processes can be seen in a scientific essentialist light, whereby the less general laws of motion are essential properties of the process natural kinds significantly lower down the hierarchy than the PLA. Neither of these explanations explains the PLA, but this is not problematic here, as we are trying to give an account that captures the more traditional laws of motion.

8.8.2 The Best Necessitarian Account of the PLA

Just as the best necessitarian account of the 'laws of motion formalism' requires no direct reference to the PLA, so I believe the best metaphysical account of the PLA formalism requires no reference to the dispositional properties inherent in the dispositionalist account of the laws of motion. It can be cashed out entirely in terms of the essential properties of the global kind.

If we are to know how the actual world is to evolve, it is not sufficient to say that the PLA is one of its essential properties. The mathematical formalism requires not just the PLA, but also the Lagrangian for the system itself. Just as Newtonian Mechanics and General Relativity have their own Lagrangians, so will the ultimate physical theory - the grand unified theory of quantum gravity. Although this is unknown to us, it is a property that, in Ellis's words, is a real essential property out there for us to discover. If we now take the Lagrangian of the unified theory of quantum gravity and the PLA to be the two essential properties of our physical system, then the entire evolution of the physical system between any two points in its velocity-configuration space can be derived.

The neo-dispositionalist and scientific essentialist interpretations of the individual laws of motion cannot, on the face of it, explain the PLA. Neither can the metaphysical interpretations of the PLA be used to explain lower-order laws of motion (or at least they are not the *best* explanations) – so how could we choose one metaphysical explanation over the other? It is not as obvious to me as it is to Katzav that we should immediately assign preference to the PLA as the best explanation of the evolution of our physical system, as the dispositionalist account seems to have the advantage of explaining events at a much smaller scale, and in the way we tend to interpret change in the world. Similarly, it is not obvious that we should prefer the neo-dispositionalist or lower-order natural kinds account, as the PLA version is, on the face of it, significantly more parsimonious. We seem to be left either with causal overdetermination, where two distinct kinds of metaphysical glue determine the evolution of the system, or with metaphysical underdetermination, where only one theory is the right one, but there is no possible way to choose between them[93].

8.9 Conclusions

The aim of this chapter was to see if applying the three standing metaphysical theories to (possibly) our most fundamental physical principle could help us eliminate one or more of the candidates. However it seems firstly that (in their own ways) all three remaining candidates (the Armstrongian view having been ruled out) can cope with the challenges provided.

However, Ellis has both dispositional properties and essential properties of natural kinds in his ontology, and therefore Scientific Essentialism does seem to be the most compatible with my claims in 8.8. But if asked "what explains the evolution of our physical system", one has to wonder which metaphysical explanation Ellis would give. Neo-dispositionalism, too, is faced with a problem: he has to suddenly allow for a major cosmic coincidence – that all particulars omnispatially and

[93] At least not without engaging in some serious metametaphysics, which is way beyond the scope of this thesis.

omnitemporally instantiate the PLA disposition. Humeanism, then, looks to come out on top.

Chapter 9: The Case Against Dispositional Essentialism

In this chapter I will show why I think we should reject dispositional essentialism, and hence both neo-dispositionalism and scientific essentialism. I will first relate Barker's argument concerning the categorical nature of the SM-relation, to Bird's 'Ultimate' argument against the contingent necessitation view of laws, demonstrating how Bird inadvertently throws his own metaphysic into the flames; secondly I demonstrate that neo-dispositionalism is simply a *variant*[94] of Armstrong's metaphysic; thirdly I show that dispositional essentialism cannot deal with counterfactuals; and finally that the dispositional essentialist has no workable account of metaphysical necessity.

9.1.1 The Ultimate Argument Against the Neo-Dispositionalist

In this section I argue that if Bird wants to uphold his 'ultimate argument against Armstrong's contingent necessitation view of laws' (Bird 2005), then he is hoisted by his own petard. Following on from section 7.6.3 I show that, when it comes down to it, neo-dispositionalism is just another view whereby a necessitation relation (albeit one that plays a constitutive role) is supposed to determine the patterns of property instantiations, and that Bird, too, needs to show why his necessitation relation (a property of his natural properties) necessitates these regularities. Unless he takes the view that the N-relation-equivalent implicit in his ontology (the SM-relation) has its identity *primitively*, such that in all possible worlds it plays the role of determining patterns of property instantiation, he is subject to the very same regress that Armstrong is. But if Bird can claim his N-relation-equivalent has its identity as an SM-relation primitively without falling into a regress, then Armstrong can claim the same thing. But neither Bird, nor Armstrong, are in a position to make these claims.

[94] It is not the *same* view, but is similar in respects which the dispositional essentialists do not intend.

9.1.2 The Stimulus-Manifestation Relation (SM-relation) as an N-Relation

The SM-relation plays a different role from that of the N-relation of Armstrong's ontology, but only in the role it plays with respect to the constitution of natural properties. Armstrong's N-relation plays no constitutive role, and is thus entirely independent of the identity of his universals - this leaves his position obviously open to the supposedly counter-intuitive consequences of categoricalism. Bird's SM-relation is supposed to solve these problems by playing a constitutive role, fixing the identity of the properties in terms of their causal/nomological roles. The N-relation and the SM-relation *do* (or at least they are mean to) play the same role in certain other respects, however - both play the role of fixing the pattern of property-instantiations.

In the case of the neo-dispositionalist: where a dispositional property, D, is constituted by second-order relational properties, SM-relations, such that it has stimulus condition S and manifestation M, $[D_{(S,M)}]$, then whenever S and D are instantiated by particulars, and these come together in the circumstances (absence of finks and antidotes etc), M manifests (where M is partly constituted by its SM-relation to D). M manifests because of the SM-relation it bears to D – or to put it another way (in the context of neo-dispositionalism), M manifests in virtue of the SM-relation it (necessarily, due to its constitutive role) bears to D.

In the case of Armstrong's contingent necessitation view: when the first order universal, F, bears the second-order relation, N, to the universal, G, in world, *w* – that is, N(F,G) is true in *w* - whenever F is instantiated in *w*, then so is G. G is instantiated in virtue of the N-relation it bears to F, just as M is instantiated in virtue of the SM-relation it bears to the coming together of instantiations of D and S. In both cases, the relations central to their respective versions of metaphysical necessity are second-order properties, whose relata are first-order (natural) properties that can be instantiated by particulars. The SM-relation, and the N-relation are for all intents and purposes the same second-order property - they merely play a

different role in the constitution of natural properties in Bird's and Armstrong's ontologies.

9.1.3 The Ultimate Argument Against Bird's Dispositional Monist Account of Laws

The conclusion reached above, on the face of it, might not overly concern the dispositionalist. Dispositionalists are on the whole more sympathetic to Armstrong's view than they are to Humeanism, as this N-relation is supposed to provide the same kind of constraints on property instantiations as they want (where Humean worlds seemingly have no such constraints). Armstrong's metaphysic is usually thrown out due to its categoricalism (and the problems arising through quidditism), and the neo-dispositionalist just solves these problems by claiming that natural properties are actually *constituted* by the second-order relations between natural properties. So why shouldn't they just embrace the thought that: yes, SM-relations are the same thing as the N-relation, it's just that Armstrong went wrong by not appreciating the constitutive role they must play?

Bird gave us some insight into why this cannot be the case – or at least why he cannot accept this to be the case. For Armstrong, claims Bird, the N-relation is just another universal (albeit a second-order universal) with a quidditistic nature. There needs to be some *reason* why, when the N-relation holds between F and G (for any F and any G), there is a regularity between Fs and Gs. To explain why N(F,G) explains the constant conjunction of instantiations of F and instantiations of G, Armstrong needs a further explanation for the regularity 'whenever N stands between two properties, there is a constant conjunction between those properties'. As the explanation cannot be just the regularity R((N(F,G),R(F,G))[95], he requires a further, higher order necessitation relation, N'. As we saw in section 5.6 this generates a regress, as we also need an explanation for why N' necessitates the regularity between N-

[95] Where R(F,G) says, there is a constant conjunstion between Fs and Gs

relations and regularities – so we postulate N'', and so on. So, for Bird, the SM-relation in neo-dispositionalism cannot be like the N-relation in Armstrong's view - firstly because if it were, we would be introducing quidditism back into our ontology (and the very thesis of neo-dispositionalism is that it is devoid of quiddities), and secondly because, if the SM-relation were the same second-order property as the N-relation (in all respects except its constitutive role), there would be an equally vicious regress in neo-dispositionalism.

But this, I contend, is exactly what we end up with. In 7.6.2 I provided an exposition of Barker's arguments that demonstrates why the SM-relation cannot be a powerful disposition: he argued firstly that the SM-relation is a second-order property – a property of properties, whereas dispositional properties are first order natural properties of *things*; and secondly because if SM-relations are powerful, then they must have their identities fixed by third-order relations to other second-order properties, and these third-order relations in turn must have their identities fixed by fourth-order relations to other third-order properties, and so on. This regress is, as Barker points out, a vicious one.

To recap: the identity of the natural properties *can* be dispositional properties fixed relationally, but if (a) the properties of natural properties are to be powerful, and (b) the identities of natural properties are to be relationally constituted by SM-relations, then the SM-relations must be powers. But the identity of all powers needs to be relationally constituted, including, it seems, the SM-relation, the relations that constitute the SM-relation, the relations that constitute the relations that constitute the SM-relation, and so on. We simply cannot fix the identity of all the requisite relations without generating this vicious regress.

If the identity of the SM-relation cannot be fixed relationally, then it is not a power – it is a quidditistic, categorical, second-order property. The identity of dispositional, natural properties, and therefore 'what it is to possess a causal or nomological nature' (Barker 2009: 7) is thus fixed by a

pattern of categorical second-order properties. If Bird wants to argue that Armstrong's position fails in virtue of the fact that the N-relation is a quidditistic property, and attempting to fix its identity by using higher-order N-relations generates a vicious regress, then he is hoisted by his own petard, for the SM-relations that fix the identity of natural properties (and hence their causal/nomological roles) are equally quidditistic - fixing the role of the SM-relation using higher-order SM-relations generates the exact same regress. Bird's account of laws as internally governing relations between first order natural properties, seems to fare no better that Armstrong's externally governing conception.

9.1.4 A Solution?

Perhaps the neo-dispositionalist could simply say that it is a primitive matter of fact that in all metaphysically possible worlds, the SM-relation plays the role that it plays in our world; that is, it is a brute fact that it determines (with metaphysical glue) the patterns of property instantiations in all metaphysically possible worlds. But it seems to me that if Bird responds in such a way, then Bird cannot use his regress objection against Armstrong. Bird's argument against Armstrong relies on the N-relation being just a standard quidditistic (second-order) property that could change its role across possible worlds – but if Bird is allowed to make the claim that his SM-relation plays the same role in all possible worlds as a brute matter of fact, then why can't Armstrong? Armstrong could, by the same logic, simply respond that the N-relation is the primitively powerful relation that provides the proponent of the contingent necessitation view of laws both a means of judging between accidental and non-accidental regularities, and a metaphysically meaty explanation of many of the regularities in our world.

But this looks messy for both accounts: for the Armstrongian because it introduces a second kind of property – a universal that is not categorical; and for the neo-dispositionalist because he wants an ontology completely devoid of primitives of this kind (after all, it is precisely

because of similar primitives in the Humean ontology that Humeanism is so unpalatable for the dispositionalist) - but how is Bird to justify his claim that these second-order properties play the SM-relation role (as opposed to some other kind of ordering) in all metaphysically possible worlds? Perhaps the dispositionalist will just have to accept that the chain of explanation must end somewhere, and this is where it ends – the primitive powerfulness of natural properties. I find this somewhat unsatisfying, however, as we have still not really been given an account of the *nature* of dispositional properties (that is, the natural properties' properties) – which, when it comes down to it, I think the main project of the dispositional essentialists should be.

9.2.1 Dispositionalism is Just Armstrongianism in Disguise

In this section I will show, following Barker (forthcoming), why it is unsurprising that Bird's argument against Armstrong can be turned on his own position, by demonstrating that in fact, the whole dispositional essentialist project is just another form of the Armstrongian position. The laws of the neo-dispositionalist are no different from the laws of the contingent necessitation view – at least not in a way that shows it to be a new *kind* of necessitarian position.

To recapitulate Armstrong's view: laws constrain the patterns of property instantiation, and these laws are natural necessitation relations holding between universals. N(F,G), which denotes the N-relation standing between two natural properties, constrains all metaphysically possible worlds in which it holds, so that it is not possible, within those worlds, for the state of affairs Fa and ¬Ga to be instantiated. The modality (what is possible and what is not within a metaphysically possible world) in Armstrong's view is thus provided by the N-relation, but the actual holding of the N-relation is contingent in so far as there are N(F,G) worlds, and there are non-N(F,G) worlds. The N-relation is a universal, and so to be consistent with Armstrong's view of properties, it should be categorical, and thus quidditistic.

To summarise neo-dispositionalism: The neo-dispositionalist believes that all properties have a dispositional essence, and that [the identities of] these properties are relationally constituted. They can claim that there are no quidditistic natural properties, because all the identities of natural properties are fixed by their place in an asymmetric 'graph', or pattern of SM-relations (stimulus-manifestation relations). It is not, therefore, metaphysically possible for any property to play any causal/nomological role other than the role it actually plays, in this, or any other possible world. The modality in neo-dispositionalism thus flows from the dispositions of natural properties. As we saw in section 7.2: in all possible worlds, if *x* instantiates dispositional property, D, if the stimulus conditions were met, then P would manifest.

Neo-dispositionalism, therefore, supposedly differs from Armstrongianism in two important respects: firstly, the properties in Armstrong's metaphysic are all quidditistic – they can change their causal/nomological roles across possible worlds; and secondly, the necessary connections in neo-dispositionalism are meant to be found in the intrinsic (always dispositional) natural properties of the things that instantiate them, whereas in Armstrongianism the necessity is provided by a further (second-order) universal that's extrinsic to the natural property instantiations – the N-relation between first-order universals.

9.2.2 A New Armstrongian View – What *Constitutes* A Dispositional Property

According to Armstrong, the laws are contingent as N(F,G) holding in any world is a contingent matter of fact. But Barker demonstrates an alternative Armstrongian view, Armstrongianism*, wherein the main principles of the Armstrongian view of laws are upheld - that is, the properties are quidditistic, and the N-relation provides the metaphysical glue – but the laws of nature are nonetheless metaphysically necessary. Armstrongianism* can be summarised as follows:

1. All properties are quidditistic – they have their identity primitively, not being fixed in any way by their modal character.

2. The N-relation ensures that whenever N(F,G) and Fa, then Ga

3. Properties have their role '*insofar* as' they instantiate certain N-relations to other quidditistic properties.

4. Each natural, quidditistic property has a unique position in a network of N-relations.

To clarify:

> The *insofar* as-locution is a way of augmenting the identity conditions for some object. For example, a person is, roughly, a passenger insofar as they are conveyed by some vehicle for some journey, which means the identity conditions of the passenger *x* is that they are identical to person *O* and that *O* is associated in the right way with some journey-vehicle pair. Applying that idea to the term *p insofar as it has network-position-Y*, the latter denotes the entity *x* that is (i) identical to quiddity *p*, and (ii) in the network position *Y*. (Barker: forthcoming)

The upshot of this is that the laws of nature are metaphysically necessary. The natural property, charge -1, could not play a different role – that is, it could not abide by different laws – as *x* is a natural property insofar as it fits into its particular position in the network of N-relations. This is not merely a *de dicto* necessity[96]. Take a 'phase' to be a section of the *x*'s existence across modal space - just as a person can be a traveller at one time and not at another, so a property can fit into the network of N-relations in the right way at one world, and not at another. As the natural property is the quiddity when it is in certain 'phases', and these phases are real and not linguistic constructions, these phases corresponding to the natural property (the ones where the property has the right position in the network of N-relations) can provide the foundation of the *de re* necessity of Armstrongian* laws. As Barker puts it: 'the nature of the [natural]

[96] One might assume that there nothing to a natural property but the quiddity, but this would be false – the phases are real.

property is made up of a quidditistic core, but overlaid by modal relations – it's a quiddity under certain conditions' (*ibid*)[97].

Armstrongianism* about laws, although it is different from Armstrongianism about laws insofar as if N(F,G) holds in one world, (whereby F and G are natural properties), then N(F,G) is true in all possible worlds (or at least in those worlds in which F is instantiated), yet it remains Armstrongian in the sense that the universals are quidditistic, and the N-relation provides the modal force. Clearly this differs significantly both from the Humean position, and, on the face of it, from dispositional essentialism. However, these apparent 'significant' differences to dispositional essentialism are really just superficial.

The first reason we have for thinking this has already been shown - the second-order property linking the natural properties in Armstrongianism* is the categorical N-relation, and the second-order property linking the natural properties in neo-dispositionalism is the, what turned out to be categorical, SM-relation. The only difference was the role the SM-relation plays in the constitution of the identity of the neo-dispositionalist's natural properties. So far, though, we have been talking about the constitution of a property's *identity*, rather than what constitutes the property itself – indeed, the dispositional essentialist never provides us with such an account. The next section will address this issue.

9.2.3 What Constitutes a Dispositional Property?

There are theories about the constitution of particulars in terms of first-order properties. Some think an object is just a bundle of natural properties; others think of 'thin-particulars' upon which first-order natural properties attach themselves - but what about the constitution of natural properties in terms of second-order properties? Barker suggests there are two comparable possibilities for the neo-dispositionalist: natural properties are bundles of SM-relations (whose relata are other bundles of SM-

[97] My thanks to Stephen Barker for our discussions of this issue.

relations), or natural properties as 'thin-properties' upon which SM-relations (whose relata are other thin properties) attach themselves.

Take the second option – that natural properties are thin-properties, 'thickened' by SM-relations to other thin-properties. This immediately looks troublesome for the neo-dispositionalist, who, remember, wants an ontology devoid of properties with inherent categorical natures. The thin-property itself is, it seems, the property-equivalent of a substratum. The thin property has no modal role, as the natural property's modal role is played entirely by the SM-relations that thicken it. Properties end up having a categorical core, just as with Armstrongianism*.

In fact, what we are left with looks an awful lot like Armstrongianism*. Even if we assume, as the neo-dispositionalist must, that the thin-properties have their particular SM-relations necessarily (they have their identity fixed by their place in the network of SM-relations), this does not make the properties powerful. We have a categorical base in the form of (modally inert) thin-properties, thickened only by the (categorical) SM-relations that constitute the network of relations that fix the identity of the natural properties. No 'powers' in sight, and plenty of quiddities!

The bundle approach does no better. Relations have an order - N(F,G) entails, by Armstrongian logic, that all Fs are Gs, but not that all Gs are Fs. When a property instantiates an SM-relation, we have to know which 'docking-point'[98] the property is 'attached' to – for the neo-dispositionalist, if natural property/disposition F is partially constituted by the SM-relation connecting it to G, where G manifests when F's stimulus conditions are met, then the ordering requires F to be prior to G. A 'bundle' first-order property cannot, therefore, simply be the bundle of the second-order properties that constitute it, for this would not take into consideration the orderings. We can conclude that only the relevant docking point should be considered as in natural properties' bundles. But,

[98] Barker's terminology

this leads to all properties being numerically identical, as 'all natural properties are indiscernible in that each natural property has both docking points of [the SM-relation] in its bundles, since each natural property enters into both positions in SM-relational facts, that is facts like N(F,G)' (*ibid*).

The bundle theory for natural-property-constitution is hopeless, and the 'thickened', thin-properties just give us a position resembling 'Armstrongianism*', as the properties turn out to be quidditistic – a position the neo-dispositionalis deplores. Bird says that laws 'supervene' on dispositional properties, and that these dispositional properties can govern 'internally'. Well, it seems to me that Bird's properties, which supposedly have their 'power' through the second-order SM-relations, govern just as Armstrong's laws govern; that is, via what turned out to be categorical relations between first-order natural properties. Lewis said that being called Armstrong does not give you massive biceps – well, calling a property a power does not make it powerful, either.

9.3.1 Dispositional Essentialism and Counterfactuals

One of the main necessitarian arguments against the naive regularity theorist was that this earlier form of Humeanism could not deal with counterfactuals. David Lewis and his Humean possible worlds analysis of counterfactuals, however, deals with assessing counterfactuals very nicely. But can neo-dispositionalism do the same thing? I argue not – and furthermore, and for similar reasons, it turns out to be very difficult for the neo-dispositionalist to give any account of modality.

Dispositional essentialists no doubt think they have the upper hand on the Humean when it comes giving truth values to counterfactuals – the statement 'if I had not struck the match, it would not have lit' is made true, they would argue, by the natural, dispositional properties instantiated by the match and by the match-box. There is a *reason* why the match lit – namely that the stimulus conditions required for the requisite dispositions to manifest were obtained. But as we shall see, it turns out that

dispositional essentialists have a serious problem when it comes to counterfactuals.

9.3.2 The Problem of Counterfactuals

Let us assume possible world semantics in a deterministic world, and consider the example we have above. As we have seen, with the Humean conception, if the match *was* struck in the actual world at time, *t*, (and it lit), the nearest possible world in which that match was not struck at *t* will require a small, localised miracle (with respect to the laws of the actual world); that is, if we take the laws of our world, and the precise match of matters of fact between the actual world, *w*, and the close possible world, *w**, prior to *t*, then assuming determinism, the match would *have* to be struck and lit in *w** at *t*, unless we allow for the laws to be breached. The Humean thus allows for the small, localised miracle, and then lets *w** run on in accordance with the laws of *our* world (as opposed to the laws of *w**, which would almost inevitably differ from the laws of the actual world)[99]. We then judge the truth value of the counterfactual by looking at the states of affairs at *w** after *t*.

But this cannot be the case with dispositional essentialism. For the dispositional essentialist, the laws are necessary – the laws supervene upon properties, and thus the laws are identical in all worlds with the same properties. By the very nature of the neo-dispositional essentialist's 'modality', there are no metaphysically possible worlds with minor miracles of the kind envisaged in Lewis's theory. It turns out that a world with a course of events similar to that of our world up to time *t*, but without the match's striking at *t*, must be a world with very different properties, and thus very different laws. But if the laws in the closest 'non-striking' world are very different from the laws of the actual world, then there is no reason to suppose that, in such a world, a match that is not struck fails to light. It follows that the neo-dispositionalist has no

[99] as opposed to the laws of *w**, which will differ from those of *w* (as the best systems analysis does not allow for laws (regularities) with exceptions).

justification for asserting his counterfactual: “if the match had not been struck, it would not have lit”.

Perhaps one might argue that we can imagine a world, *w***, that comes into existence in a state that matches our world at *t* perfectly, except in the respect of the match, which is not being struck at *w***. We can let *w*** run on in accordance with the laws that supervene on the world’s properties (the same properties, and thus same laws as the actual world), and judge the truth value of the counterfactual. But we end up with a situation where the whole histories of *w* and *w** would bear very little resemblance to one another. The closest non-striking world would be a world very different in its history to the actual world, and this would give counter-intuitive results when assessing counterfactuals in accordance with the possible world semantics.

9.4.1 Metaphysical Necessity in Neo-Dispositionalism – A Worry

In section 7.2 I gave the account of modality that the neo-dispositionalist would wish to uphold. Given that there are no finks or antidotes at the fundamental level, in a deterministic world it is *de re* necessary that if the (dispositional) natural property, D, is instantiated by *x*, and that the stimulus conditions for D are met, then D’s manifestation property will be instantiated (possibly by *x*, but not necessarily – it depends on the nature of the manifestation property); in other words, in all possible worlds in which D is instantiated by one or more particulars, and the stimulus conditions for these instantiations of D are met, the manifestation property or properties manifest in accordance with D’s place in the network of SM-relations. But, I claim, the possible worlds account of modality fails when applied to neo-dispositionalism, and if the neo-dispositional essentialist is to have any account of modality at all, he will have to devise an entirely new one.

There are two ways of approaching the possible worlds account of modality, the modal realist account, and the ‘worlds as abstract entities’ approach. The modal realist approach regards every possible world as a

real, existing 'universe' (perhaps one might call them the 'real, existing worlds that constitute the multiverse). A proposition is necessarily true if it is true at all of these possible worlds - electrons have charge -1 necessarily if all electrons at all possible worlds have charge -1, and so on. A proposition is contingently true if it is false at any possible world - 'There are stars' is contingently true, because although there are stars at our world, there are no stars at many others. Another 'possible worlds' view of modality takes possible worlds to be abstract entities, whereby there are no 'real' worlds other than the actual world. The proponent of this view believes we can judge our modal claims through this kind of abstract entity – often envisaged as a set of maximally consistent propositions. Take the abstract world, *a*, corresponding to the actual world, *w*; that is, the abstract entity including an abstract form of every state of affairs in the actual world (for a particular giraffe, *g*, in *w*, there is an abstract entity *g** in *a*). In the following sections I show why neither of these views of possible worlds seem to allow for a dispositional essentialist account of modality.

9.4.2 What determines the property instantiations – Questioning metaphysical necessity in Dispositional Essentialism?

By now one is used to talking about 'metaphysical necessity' in the necessitarian metaphysical accounts of law and causation. An event is called metaphysically necessary when some kind of modal-force, or necessary connection links two properties together. It is the N-relation in Armstrong, and the equivalent second-order properties of dispositions in dispositional essentialism. In this section, though, I question whether dispositional essentialists have the right to claim that the nature of natural properties metaphysically necessitates the pattern of events in our world.

The notation of N(F,G) in Armstrongian accounts of modality is familiar, but there is no reason why we cannot use the same, or at least similar notation for dispositional essentialist accounts, too. Call this N(F,G)*. F and G, for the dispositional essentialist, are natural properties

in the form of universals. N, when applied to F and G to give us N(F,G)*, denotes an SM-relation holding between F and G – where F is the stimulus for the manifestation property, G (the ordering tells us which is the stimulus and which is the manifestation). There is, of course, a difference between Armstrong's N(F,G) and N(F,G)*, in that N plays a constitutive role in the latter – nonetheless, N(F,G)* tells us all we need to know in this context.

According to the dispositional essentialist, if N(F,G)* is true, then all Fs are Gs (or in some cases, all Fs are followed by Gs, depending on what F and G stand for – let us stick with all Fs are Gs for simplicity) in all possible worlds. So what does the constraining? What makes it the case that all Fs are Gs? Well, for the dispositional essentialist, the answer to this question must be 'N' – that is, the second order relation that is a property of both F and G. There can be no F that is not a G, and this is made true by this 'constraining' second order property, the SM-relation, that F and G stand in to one another – it is a part of these properties' nature. So how can this be demonstrated in terms of a possible worlds account of modality?

Firstly, consider the Armstrongian N(F,G), again. Given that N(F,G) holds in a world, *w*, it is 'contingently necessary' that all Fs are Gs. 'N' does the constraining such that it is not possible in an N(F,G) world that there is an F that is not a G. But what if N(F,G) does not hold? There can be two (what happen to be) qualitatively identical worlds, where the patterns of first-order property instantiation are identical, but in one world N(F,G) holds and in the other it does not. In the second world it is a contingent fact that all Fs are Gs, but in the first world it was necessary given N(F,G). There is an asymmetry here that allows Armstrong to point to the N-relation as the relational property present and constraining the property instantiations in the first world. However, the dispositional essentialist cannot make such a claim, as there are no examples of an asymmetry of this kind – whether the natural property-instantiations are constrained by N(F,G)*, or whether N(F,G)* is

constrained by the natural property instantiations is not demonstrable. The dispositional essentialists just *tell* us that the former is true, without giving us sufficient reason for thinking so. N(F,G)* could just be a fact made true by the pattern of property instantiations in the metaphysically possible worlds.

It looks pretty implausible that these doubts should arise, but even a brief look at the metaphysically possible worlds in a dispositional essentialist metaphysic, will show that these doubts are fully justified. The dispositional essentialist's N-relations, or SM-relations as I called them before this comparison with Armstrong, constitute the identity of a property. So if F and G are instantiated in a world, then so is N(F,G)*. It is a simple consequence of the dispositional essentialist metaphysic that here are no possible worlds where 'F and G and not-N(F,G)*', so, it seems, there is no way of judging what is constraining what.

The dispositional essentialist might respond by saying that N(F,G)* constrains what possible worlds there are. This cannot be right. Possible worlds just *are*, or they *are not*. Properties are not, it seems to me, prior to worlds in such a way that they can determine which worlds come into existence and which do not. Which possible worlds exist and which do not is not determined by some kind of sieve, filtering out the bad ones. Worlds, at least for the modal realist, are just there in the form of four-dimensional blocks of property instantiations. In desperation the dispositionalist might appeal to the truth-values of counterfactuals, but as I have already shown, dispositional essentialism does not give us a plausible means of judging counterfactuals, so this option is not available.

The dispositional essentialist can therefore give us no reason to think N(F,G)* constrains the property instantiations rather than vice versa. They *can* say that N(F,G) constitutes F and G as instantiated in our world, and also that it constrains them in this world, but only if we take constraint to be a special kind of regularity relation – one that holds between natural properties and the properties of those natural properties (a constant

conjunction), and with a primitive 'direction' (the second-order properties primitively constrain the first-order). But this claim has nothing interestingly modal in it at all – we have just defined constraint in such a way that it conforms to the usage the dispositionalist wishes to impose. As it stands this does not capture our conception of constraint (in terms of metaphysical glue) at all, so the notion of 'internally governing properties', (or in other words, 'the nature (second-order properties) of natural properties constraining property instantiation') which Bird proposes, has little justified metaphysical substance. At this time I will not be so bold as to say dispositional essentialists cannot come up with a modal picture that does, in a more metaphysically meaty sense, capture our concept of 'governing intrinsic properties', but at the very least I hope I have shown there is more work to be done.

9.5 Chapter 9 Conclusions

In this chapter I have tried to provide good reasons for rejecting the dispositional essentialist metaphysic. These are serious challenges the dispositional essentialist has not yet met, that even question whether dispositional essentialism really is a new way of looking at laws and causation; it turns out, I have argued, that dispositional essentialism is simply an Armstrongian-style view, but one where the dispositionalist's SM-relation plays a constitutive role.

As a result of this conclusion, it should hardly be surprising that Bird's argument against Armstrong proves seriously problematic for his own metaphysic. Of course, these problems arise from claims about the nature of dispositional properties that Bird would deny - namely that the identity of dispositional properties is ultimately constituted by second-order *categorical* relations, with a quidditistic base. A satisfactory answer to *how* natural properties are relationally constituted, as opposed to how their identities are relationally constituted, might have solved this problem. But the dispositionalist never provides us with an answer to this question, and with good reason: because a satisfactory answer cannot be

found! Both of the two possible answers, namely the bundle theory and thin-property theory, fail to produce a successful account of property constitution.

Finally, I questioned the dispositional essentialist's account of modality; that is, how he explains the metaphysical necessity in his view of laws and causation. Armstrong, it seemed, had a reasonable way of demonstrating that the N-relation constrained the property instantiations and not the other way around (as his necessitation relation holds contingently at worlds). However, the dispositional essentialist's account is not asymmetric in this way - the dispositionalist never provides us with an account of what relational constitution actually is, and furthemore they have no way of showing what constrains what – the neo-dispositionalist simply tell us what way they want it to be.

For the four reasons I have articulated in this chapter, I conclude that as things stand, we have no reason to accept dispositional essentialist views of any kind as plausible metaphysical accounts of cause and law.

Chapter 10: Thesis Conclusions

My aim in this thesis was not to devise a perfect metaphysics of cause and law. Nor was it fully to endorse any individual existing metaphysic. My aims were far more modest: to discover the most coherent account of our conception of causation, of our conception of laws, and the link between the two; to find out what conditions metaphysical accounts of laws and causation must satisfy; and to examine existing (and conflicting) metaphysics to determine their consistency with our concepts, their ability to stand up to the main objections made against them, and to see whether these metaphysics achieve what they set out to do. The conclusion of this thesis is to reject the existing anti-Humean metaphysics. There were three main anti-Humean options: Armstrongianism, Scientific Essentialism, and Dispositional Essentialism. All three are untenable accounts.

10.1 Our Concept of Causation

The opening chapters discussed the nature of our conception of cause and law. This, I believed, was necessary, as were we to start presenting metaphysical accounts of cause and law that did not cohere with our conception, as a matter of *de dicto* necessity we would not be talking about causation and laws of nature at all. I concluded that the most coherent and intuitive account requires causation to be analysed in terms of counterfactuals. Although temporal priority and spatial contiguity were ruled out as necessary conditions, its asymmetry, and the link between cause and law were deemed fundamental aspects of our conception of causation.

Conceptions of laws of nature can differ whilst remaining internally consistent – what I wished to unravel were those aspects of this conception common to all consistent conceptual analyses. Universality, at the fundamental level at least, I deemed to be a necessary condition of a law of nature. *Ceteris Paribus* laws are often discussed, but it seems to me

that violations of these laws are not violations at all. When I fall from a plane and do not continue to accelerate at 9.81 metres per second per second, due to air resistance, this is not a violation of the law of gravity. Laws constantly 'contribute' the same to each motion, and this can be accommodated in all metaphysical analyses. Another suggestion was that laws must govern, but many have perfectly consistent non-governing intuitions about laws. 'Governing' cannot, therefore, be a necessary condition of our concept of a law of nature.

It was later shown that our concept of cause and our concept of laws of nature share many features – in particular, they are both appealed to when explaining and predicting events. The question "why did I fall to the ground?" can be answered both in terms of cause and in terms of law: "because David pushed me", and "because of the laws of gravitation" are both reasonable responses; which explanation we choose merely depends on the context in which the question is asked. It was unsurprising, therefore, to find out that in at least three of the four metaphysics of cause and law considered, either the account of the former featured prominently in the account of the latter, or vice versa: In Hume's regularity theory and in Armstrong's governing conception of laws, an instance of causation is simply an instance of law; in Lewis's more sophisticated, Best Systems Analysis of Laws, causation is analysed in terms of counterfactuals – but the truth values of the counterfactuals is determined by the laws of nature; and in Bird's dispositional essentialism, the laws supervene on the dispositional properties of things, that is, the underlying metaphysics of causation, and the underlying metaphysics of laws, are ultimately the same phenomena: dispositional properties.

Ultimately, regardless of the metaphysics underlying cause and law, my conceptual analysis lead me to these two most important conclusions:

1. The cause-effect relationship, from a conceptual analysis perspective, should be analysed in terms of counterfactuals.

2. The propositions *expressing* the most fundamental laws of nature are universal statements, that, together allow us to explain and predict the evolution of the entire physical system, in the simplest and strongest possible way.

10.2 The Metaphysics of Causation

Many metaphysicians believe that the suggestion that laws (and *ipso facto* causes, given the metaphysics of the relevant philosphers) supervene on the evolution of the physical system and not vice versa is entirely unacceptable. This, they claim, deprives them of any ability to provide explanations of events, and any right for us to use them to predict future events or make inductive inferences about future states of affairs. As a result of these beliefs, anti-Humeans have devised a number of possible alternatives, the major contestants being the Armstrongian governing conception of laws, Scientific Essentialism, and Dispositional Essentialism. Not only did I contest their claims that Humeanism fails in the above respects, but I also showed that all three anti-Humean accounts fail by their own standards when we consider the implications of their various metaphysics.

In chapter 8 we saw that dispositional essentialism would, in principle, be able to account for possibly our most fundamental scientific principle – the Principle of Least Action. But this required the prima facie extremely unlikely event of all particulars omnitemporally instantiating the PLA disposition. However, dispositional essentialism, I argued, is not tenable as an account of causation - firstly because it cannot account for counterfactuals, which form the basis of our concept of causation (and so, as a matter of *de dicto* necessity, theirs is not an account of causation); secondly because it is subject to an identity regress (with the failure to fix the identity of the SM-relation); thirdly because they cannot provide an account of the constitution of dispositional properties in terms of their second order properties (SM-relations between natural properties); and finally it fails because the dispositional essentialist does not achieve what

he sets out to - namely to find a completely different metaphysical approach to cause and law to Humeanism and Armstrongianism, that also provides us with metaphysical glue. Dispositional Essentialism, it turned out, collapses into an alternative form of Armstrongianism, an account that Bird himself rejects by demonstrating how it falls into an identity regress. Furthermore, the SM-relations, that is, the second-order properties of natural properties for the dispositional essentialist, turn out to be categorical. If the properties of natural properties are all quiddditistic, the dispositional essentialists are subject to the very same 'quidditism' objections they raise against both the Humeans and Armstrong[100]. Dispositional Essentialism is not, therefore, a viable anti-Humean metaphysic.

Scientific Essentialism fails for similar reasons: although Ellis's account of laws is based upon the essential properties of natural kinds, the evolution of the physical system, for Ellis, depends upon dispositional properties as well as categorical properties. There are many reasons, I think, to prefer scientific essentialism to dispositional essentialism, but the mere existence of properties that are wholly dispositional leads to the same problems faced by the neo-dispositionalist – without powerful properties in Ellis's ontology, there is no oomph.

All forms of Armstrongianism fail. Armstrong claims that his natural necessitation relation 'contingently necessitates' the events/property-constant-conjunctions it stands between. However, there is little in the way of metaphysical explanation to substantiate this claim. Armstrong's N-relation is a universal (albeit a second-order universal) just like any other, and as such it is a categorical, quiddditistic property. The N-relation plays an identifiable role, but if it is quiddditistic, there needs to be some explanation for why this relation always plays the role that it does. If this explanation is a regularity, then Armstrong's own position would (a)

[100] I reject the claim that these quidditism objections are problematic over and above their perceived counter-intuitiveness, but the ability of dispositional essentialism to avoid the quidditism objections is one of the biggest motivators for their position, so I deem it a strong objection to dispositional essentialism nonetheless.

force him to reject the explanation as viable, and (b) negate the need for the N-relation in the first place, as the regularity would suffice as the best explanation at the first-order level. On the other hand, if Armstrong posits the explanation to be a third order N-relation we end up with a regress, as this third order N-relation holding in all worlds itself requires an explanation. As we have seen, the regress formed is vicious, and to all intents and purposes rules Armstrongianism out as a contender. Humeans, it seems, are the last men standing.

In chapters 3 and 4 I discussed two forms of Humeanism: Hume's naïve regularity theory of laws of nature, and Lewis's sophisticated version, the Best System Analysis. Lewis followed Hume in postulating both that laws are universally quantified statements, and proposing that there is no metaphysical necessity in the world – there is no metaphysical glue tying events to one another, such that an 'effect' event occurs with metaphysical necessity given the occurrence of a 'cause' event; that is, in a world, *w*, where event X causes event Y, there are many possible worlds in which the events preceding and including event X match *w* perfectly, but event Y does not occur. The main objections to Hume's account were as follows:

1. Humean laws do not explain events, as a regularity cannot explain one (or more) of its instances;

2. Humeans do not have the right to reason inductively, so we cannot use them to predict future events;

3. Humean laws do not govern their instances;

4. Humean laws, as mere regularities, cannot distinguish between accidental and non-accidental regularities.

5. Humeanism is a quiddditistic ontology

In chapter 3 I demonstrated that the anti-Humeans are not able to criticise Humeans on the grounds of objections 1 and 2, as it can clearly be

seen that Humean regularities, and indeed regularities in general, can and often *do* have explanatory value. In fact, to deny this would lead to absurdity even on their part, for many (explicable) every-day events can never be explained in terms of 'metaphysical glue'.

Furthermore it was demonstrated that the regularity theorist has at least as much right to reason inductively as the necessitarian. In section 3.3 I made the narrower claim that Humean had as least as much right to reason inductively as Armstrong, but as it has been shown that dispositionalist accounts collapse into Armstrongianism, this can be extended to necessitarian metaphysical positions in general.

Objection 3 above might have been successful, were it the case that laws *had* to govern their instances. But it was deemed that a coherent conception of laws could be found where laws do not play a governing role, so objection 3 holds no weight whatsoever.

Objection 4, however, is a genuine problem for Hume's naïve regularity theory. With the only requirement for a law being that it is a true universally quantified statement, there would be no way of distinguishing merely accidental regularities from non-accidental, law-like regularities - any proposition expressing a (true) single-case uniformity would be a law of nature. This, of course, is unacceptable. But this objection did not rule out all Humean positions. Lewis's best system analysis, I believe, dealt with this objection perfectly adequately.

Lewis claimed that laws of nature are universally quantified statements that describe the regularities in our physical system with the best balance between simplicity and strength; the laws are universally quantified statements that maximise the number of truths derivable from as few statements as possible. Under this account, single-case uniformities, and indeed any universally quantified statements we would consider to be accidental, would not come out as laws of nature unless nature was particularly unkind to us (or the world was a very simple one, in the case of a single-case uniformity). Lewis's account, it seems to me, would more

often than not provide us with exactly the right laws. It was easy to see when we looked at the Principle of Least Action, for example, that Lewisian Humeanism would very easily pick out the PLA as a fundamental law of nature.

Objection 5, the problem of quidditism, demonstrated some supposedly counterintuitive consequences of Lewis's nomilalist account - the most prominent of which being that the same property would play many different roles in different possible worlds. But counterintuitiveness cannot be considered a knock down objection to any metaphysical account. Bird's attempted to provide a more substantial objection, but as we saw in section 4.6.3, this attempt fails, as Humeans take *de dicto* necessities seriously. Lewis's account is not perfect. Lewis gives us no account of why the sparse properties 'glow in the dark' – there is nothing about their nature that distinguishes them from the abundant properties. I suggested we follow the line of trans-world objective similarity, but this was itself unsatisfying in many respects.

There is, I acknowledge, more work to be done before we can fully embrace a metaphysical account of laws and causation, but it seems to me that this account will, in one way or another, be broadly Humean.

Bibliography

Alexander, S. 1920 *Space, Time, and Deity*. 2 vols. London: Macmillan.

Armstrong, D. M. 1983 *What is a Law of Nature* Cambridge University Press

- 2007b *A World of States of Affairs* Cambridge University Press

Ayer, A. J. 1980 *Hume* Oxford University Press

Barker, S. 2009 Dispositional Monism, Relational Constitution and Quiddities *Analysis* vol. 69 no.2 pp.1-8

Basson, A,H 1958 *David Hume* Middlesex: Penguin Books

Beebee, H. 2000 The non-governing conception of laws of nature *Philosophy and Phenomenological Research* vol 61, no (November), pp571-594

- 2006 *Hume on Causation* Routledge

Bigelow, J 1988 *The Reality of Numbers* Oxford University Press

Bigelow, J; Ellis, B; Lierse, C 1992 The World as One of a Kind: Natural Necessity and Laws of Nature *British Journal of Philosophy of Science* 43 371-388

Bird, A. 2005 The Ultimate Argument against Armstrong's Contingent Necessitation View of Laws *Analysis* vol. 65, no 2. (April) pp147-155

- 2007a The Regress of Pure Powers *Philosophical Quarterly* vol.57 pp513-534
- 2007 *Nature's Metaphysics* Oxford Clarendon Press
- 2009 'Structural Properties Revisited' in Handfield, T. *Dispositions and Causes* (2009) Oxford University Press

Black, R 2000 Against Quidditism *Australian Journal of Philosophy*, vol 8: 87-104

Blackburn, S 1984 *Spreading the Word* Oxford University Press

Bostock, S 2008 In Defence of Pan-Dispositionalism *Int Ontology Metaphysics* 9: 139-157

Broad, C.D. 1935 Mechanical and Teleological Causation *Proceedings of the Aristotelian Society*

Brown, M. B. 1987 'Review of Stove 1986' *History and Philosophy of Logic* **8**, 116-120

Campbell, S and Franklin, J 2004 Randomness and the Justification of Induction *Synthese* **138,** 79-99

Cartwright, N 1995 *Nature's Capacities and the Measurements*: Oxford Clarendon Press

- 1999 *The Dappled World* Cambridge University Press

Castell, P. 1998 *A* Consistent Restriction of the Principle of Indifference, *British Journal of The Philosophy of Science*. **49**, 387-395

Chakravartty, A. 2003 The Dispositional Essentialist View of Properties and Laws *International Journal of Philosophical Studies* vol. 11 (4), 393-413

- 2005 Causal Realism: Events and Processes *Erkenntnis* 63: 7-31

Ellis, B 2000 Causal Laws and Singular Causation *Philosophy and Phenomenological Research* (September) Vol LXI No.2

- 2001 *Scientific Essentialism*. Cambridge Studies in Philosophy
- 2002 *The Philosophy of Nature: A Guide to the New Essentialism*, Chesham: Acumen
- 2005 Katzav on the Limitations of Dispositionalism. *Analysis* 65: 90–92

- 2009 Causal Powers and Categorical Properties *Metaphysics of Science Conference, Melbourne* http://philsci-archive.pitt.edu/archive/00004749/
- 2009 *The Metaphysics of Scientific Realism* Acumen

Everitt, N 1991 Strawson on Laws and Regularities *Analysis* Vol 51 No4 pp206-208

Goodman, N. 1983 *Fact, Fiction and Forecast* Harvard University Press

Handfield, T 2009 *Dispositions and Causes* Oxford University Press

Heil, J. 2005 Kinds and Essences *Ratio* XVIII (December) 405-419

Hempel, C 1965 *Aspects of Scientific Explanation* Englewood Cliffs

Hume, D 1739-40 *A Treatise of Human Nature* (1978) Oxford Clarendon Press
- 1748 *An Enquiry Concerning Human Understanding* (1975) Oxford Clarendon Press

Heathecote, A and Armstrong, D.M 1991 Causes and Laws *Nous* vol 25 no.1 (March) pp63-73

Holton, R 1999 Dispositions All The Way Round *Analysis* vol.59 no.1 pp9-14

Indurkhya, B 1990 'Some Remarks on the Rationality of Induction' *Synthese* **85**(1), 95-114

Katzav, J 2005 Dispositionalism and the principle of least action *Analysis* 64: 206-14

- 2005a Ellis on the limitations of dispositionalism, *Analysis* 65 (1): 292-294)

Kim, J 1973 Causes and Counterfactuals *The Journal of Philosophy* Vol 70, No 17 570-572

- 1973a Causation, Nomic Subsumption, and the Concept of Event *The Journal of Philosophy* Vol 70, No. 8 217-236

Kripke, S 1972. 'Naming and Necessity'. In Davidson, Donald and Harman, Gilbert, eds., *Semantics of Natural Language*. Dordrecht: Reidel: 253-355, 763-769.

Lewis, D 1973 Causation *Journal of Philosophy* 70 : 556-67

- 1973a *Counterfactuals*, Oxford: Blackwell

- 1976 The Paradoxes of Time Travel *American Philosophical Quarterly* 13: 145-52
- 1983 New Work for a Theory of Universals *Australian Journal of Philosophy* 61
- 1986 *Philosophical Papers* Introduction Oxford University Press
- *On The Plurality of Worlds*, Blackwell Publishing 1986a
- 1994 Humnean Supervenience Debugged *Mind* vol 103, no 412 (October) pp473-490
- 2001 Ramseyan Humility *The Canbera Plan* (In David Braddon-Mitchell & Robert Nola (eds.), *Conceptual Analysis and Philosophical Naturalism*. Mit Press. 2009)

Lipton, P 2004 *Inference to the Best Explanation* Routledge

Livanios, V 2008 Bird and the Dispositional Essentialist Account of Spatiotemporal Relations *Journal of General Philosophy of Science* vol 39: 383-395

Mackie, J. L. 1974 *The Cement of The Universe* Oxford University Press

Marmodoro, A. 2010 *Metaphysics of Powers, The: Their Grounding and their Manifestations*. Taylor & Francis

Menand, L. 2001 *The Metaphysical Club* New York: Farrar, Straus and Giroux

Mill, J.S. 1911 *A System of Logic* London: Longmans Green & Co.p 207

Molnar, G 1999 Are Dispositions Reducible *The Philosophical Quarterly* vol.49 no.194 pp.1-17

Morreall, J 1982 Hume's Missing Shade of Blue *Philosophy and Phenomenological Research* Vol 32 no.3 (March) 407-415

Mumford, S 2004 *Laws in Nature* Routledge

- 2009 Passing Powers Around *The Monist* Special Issue

Psillos, S 2002 *Causation and Explanation* Acumen

Nolan, D 2005 *David Lewis* Acumen

Noonan, H. 2007 *Hume* Oneworld Publications

- 2010 Bird Against The Humeans *Ratio* volume 23, issue 1, pp73-86

Papineau, D. 1985 Causal Asymmetry *British Journal of the Philosophy of Science* Vol 36 no.3 (September) pp 273-289

Popper, K. 1959 *The Logic of Scientific Discovery* revd. Imp. (1980), London, Hutchinson

- 1969 *Conjectures and Refutations* London: Routledge and Kegan Paul

Rollin, B. 1971 Hume's Blue Patch and the Mind's Creativity *Journal of the History of Ideas* 32, 121-127

Sanford, D. 1976 The Direction of Causation and the Direction of Conditionship. *The Journal of Philosophy* Vol 73 no.8 (February) pp193-207

Schrenk, M. The Powelessness of Necessity (Forthcoming in *Nous*)

Smart, B.T.H (forthcoming 2011) Is the Humean Defeated by Induction? *Philosophical Studies* (currently online: http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s11098-011-9767-5.)

Stove, D. 1986 *The Rationality of Induction* Oxford Clarendon Press

Strawson, G. 1989 *The Secret Connexion* Oxford Clarendon Press

- 2008 *Real Materialism and Other Essays* Oxford Clarendon Press

Suchting, W. A. 1968 Functional Laws and The Regularity Theory. *Analysis* 29 (December):50-51.

Tooley, M. 1987 *Causation: A Realist Approach* (2001) Oxford University Press

Tugby, M. 2010 The Problem of Geometrical Properties for Powers Theory *Metaphysics of Science Conference, Bristol*

Van Fraassen, B 1989 *Laws and Symmetry* Oxford University Press