DISEASE KINDS AND FUNCTIONAL EXPLANATIONS

Thesis submitted by

Stefan Dragulinescu

In Partial Fulfilment of the Requirements

for the Degree of

Doctor of Philosophy

Department of Philosophy

Lancaster University

August 2010

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ABSTRACT

The present thesis is concerned with the character of kinds in human somatic pathology and the relation that these kinds and their members have with function-based explanations. More precisely, in the first part of the thesis I investigate whether diseased organisms, grouped together on grounds of their shared pathological features, could form *natural* kinds, taking into account that the paradigmatic natural kinds are the kinds of the exact sciences. The second part of the thesis has as a backdrop the Humean/anti-Humean debate over causation (and the specific construal of explanations according to which to explain is to pinpoint causes). In this backdrop, I enquire into what sort of function-based explanations we could provide for the symptoms and pathological behaviours exhibited by diseased organisms, if we construe such organisms as members of natural kinds.

I argue in the first part of the thesis that from a metaphysical point of view, the organisms dealt with in somatic medicine form natural kinds in the same sense in which we take the kinds dealt with in the exact sciences as natural. By comparing a 'classical', exact science kind with a kind of disease, I show that whatever features are associated with natural kind membership (e.g. involvement in laws or inductions, explanatory relevance, possession of 'essential' properties, instantiation of substantive universals, etc.) there is no 'ontological gap' between disease kinds and the kinds in the exact sciences.

The conclusion that diseases are natural kinds has a certain proviso regarding the question of whether the identity of the individual members of natural kinds is dependent upon their kind membership. Should diseases not be natural kinds, the proviso says, it would be because the properties characteristic of natural kinds must have an identity-influence over the kind members. I present in addition serious problems posed by outlining such identity bearing properties.

In the second part of the thesis, I argue that function based explanations concerned with diseased organisms - if we construe such organisms as being members of natural kinds - should illuminate positive causes for the symptoms and pathological behaviours they exhibit. We could obtain such function-based explanations, I suggest, if we interpret the functioning of biological items as the manifesting of causal powers. Against the background of the Humean vs. anti-Humean debate on causation, I show that Nancy Cartwright's capacities are a plausible variant for the powers at work in 'pathological' functioning.

I argue that one could track down these capacities if one viewed healthy organisms as *nomological machines*, in the sense in which Cartwright understands this notion. I also suggest that capacities are necessary in order to vindicate general and, more importantly, singular causal claims involved in medical diagnosis and hence to found satisfactory functional explanations.

TABLE OF CONTENTS

Abstract 2

Table of Contents 3

Acknowledgements 5

Table of figures 6

Introduction 8

Chapter 1. Natural Kinds 15

Introduction 15

§ 1.1 Criteria for natural kind membership 17

§1.2 Main accounts of natural kinds 30

§1.3 Kinds in medicine 39

§1.4 Realism (and Nominalism) about Kinds 46

§1.5 Carving nature at its joints 50

Chapter 2 Essences, Substantive Universals and processes 56

Introduction 56

§2. 1 Essences and essential properties – The lockean strategy 59

§2.1.1 The Lockean strategy 60

§2.1.2 Essential properties and identity 68

§2.2 Substantive Universals and requirements VIb and VII 75

§ 2.2.1 Why have substantive universals in our ontology of natural kinds? 78

§ 2.2.2 How could substantive universals justify the N&S requirement? 80

§2.2.3 Structural properties and the N&S requirement 83

§2.3 The problem of processes and criterion VIII 86

§ 2.3.1 Lowe's stance 91

§ 2.3.2 The main strategy 94

§ 2.3.3 Could Heraclitan or Spinozan kinds really be fitted into Lowe's ontology? 98

§2.3.4 Identity and the phase non-phase distinction 102

Chapter 3. Gold vs. the Graves' Disease kind 110

Introduction 110

§3.1 Induction 113

3.1.1 The causal induction requirement (criterion II) 114

3.1.2 The co-instantiation induction requirement (criterion V) 121

§3.2 The requirement of law-involvement 129

§3.3 The cluster requirement (criterion VI_a) 137

§3.4 The N&S and non-overlapping requirements (criteria VI_b and VII) 141

§3.5 The non-phase requirement (criterion VIII) 145

§3.5.1 Transmutation and the phase/non-phase distinction 149

§ 3.5.2 Identity and phases 154

§3.6 Again, on carving nature at its joints (and on criterion IV) 160

§3.7 Further remarks on medical kinds 170

Chapter 4 Functional explanations 177

Introduction 177

§4.1 Biological functions and their role in explanations 182

§4.2 Causal powers and explanations 190

§4.2.1 The Humean vs. anti-Humean dispute 191

§4.2.2 Cartwright's capacities 198

§4.2.3 Cartwright's rationale applied to medical causes 210

§4.3 Functions and capacities 213

§4.3.1 Problems of (dys-) functional explanations 214

§4.3.2 Functioning as the manifesting of capacities 224

§4.4 Capacities in medicine 233

§4.4.1 What actually happens and what could happen 234

§4.4.2 Capacities and singular causal claims for diseases 242

§4.5 Concluding remarks on functional explanations 248

Some final remarks 250

Bibliography 254

ACKNOWLEDGEMENTS

I received financial support for writing this thesis from the Arts and Humanities Research Council (fees doctoral bursary, awarded July 2007, Ref. No. 2007/135148/Lancaster University) and the British Society for the Philosophy of Science (doctoral scholarship, awarded July 2007).

I am very grateful and indebted to Rachel Cooper for her encouragement and support, as well as for innumerable comments and discussions on the content of this thesis. I would also like to thank Brian Garvey and Mihaela Vatavu for their comments on the penultimate draft of this thesis. Last but not least, I am grateful to the two referees of this thesis, Prof. Alexander Bird and Dr. Nick Unwin, for very useful suggestions and criticism.

A part of chapter 3 of the present thesis has been published as 'Diseases as Natural Kinds' in *Theoretical Medicine and Bioethics*.

This thesis is dedicated to the memory of my grandmother, Vera Dragulinescu.

TABLE OF FIGURES

Figure 1 Criteria for natural kind membership, specifying what nature the properties and the patterns they form should have 26
Figure 2 The criteria of membership in a natural kind NK for an individual a 28
Figure 3 Key accounts of natural kinds 30
Figure 4 The switch of focus 84
Figure 5 Lowe's scheme 94
Figure 6 Commonly accepted kinds as categories 105
Figure 7 Gold kind members and organisms suffering from Graves' Disease (reproduced from Wheetman, 2000) 110
Figure 8 The criteria of membership in a natural kind NK for a particular *a* 111
Figure 9 Determining properties (reproduced from Brent 2008) 117
Figure 10 Ophthalmopathy mechanism (Reproduced from Ginsberg 2003) 119
Figure 11 Immune aspects of ophthalmopathy (reproduced from Wheetman 2000) 120
Figure 12 A part of the causal mechanism inducing the production in excess of thyroid's hormones (Reproduced from Wheetman 2000) 125
Figure 13 The co-instantiation of [p1, p2,.....pn] 128
Figure 14 Diagnostic algorithm (reproduced from Wheetman 2000) 138
Figure 15 Phase and substantial changes 150
Figure 16 Ontological dependencies between criteria of natural kind membership 173
Figure 17 Features of the anti-Humean stance 193
Figure 18 Features of the Humean stance 195
Figure 19 The Humean (left) vs. anti-Humean (right) 197
Figure 20 Capacities and inference tickets 201
Figure 21 Cartwright's rationale applied in medicine 212
Figure 22 Ventricular diastole (left) and systole (right) 215
Figure 23 Aortic leakage (reproduced from Adam Health Encyclopaedia) 217
Figure 24 The Humean stance on the C2 functioning and the attending causal claims 228
Figure 25 Capacities in medicine 232
Figure 26 Renin at the renal level 235
Figure 27 The renin enzymatic pathway 236
Figure 28 The heart and the RAAS 237
Figure 29 AoI and the RAAS 239
Figure 30 What actually happens and what could happen 240
Figure 31 Cardiac ecography (reproduced from Janelle, 2003) 243
Figure 32 A causal context constituted of a background with a different array of causal factors.. 245
Figure 33 What causes what? 246

INTRODUCTION

In science as in everyday life, we operate with classifications that organise objects according to various similarities. On a metaphysical level, the classified objects are candidates for being members of natural kinds. The notion of natural kind is thus intended to delineate a metaphysical category that reflects the fact that objects resemble one another in different degrees and that some of these resemblances seem more 'profound' than others. Banknotes, galaxies, weeds and shopping bags, for instance, certainly appear less similar than, say, metals. That some of our classifications seem to be more important than others stands in need of a metaphysical explanation and natural kinds represent (the beginning of) an answer.

By appealing to a minimal construal of properties, we could say that members of natural kinds are individual objects that share certain properties and that to look for the important resemblances between individuals is to search for certain patterns of properties that should be part of an 'ideal classification'.[1] Different authors have proposed different requirements for exactly what similarities are required for natural kind membership. These requirements refer to inductions, laws, identity, universals, causation etc., and connect the natural kinds theme to various other controversial problems in metaphysics (the most important of which being the problems of essences, hierarchies, substantive universals and processes, to which I will return shortly).

The idea that things in the world can fit into an 'ideal classification' has a definite intuitive appeal. It is important to note that natural kinds have been, in fact, associated with an entire luggage of intuitions. Theorists have often affirmed that to reveal the resemblances shared by natural kind members is to 'carve nature at its joints', 'classify objects independently of our thoughts, values, conventions and language', 'provide a basic ontology', 'divide the world in a non-arbitrary way', 'illuminate what things are', 'delineate

[1] Bird and Tobin note that if there were a single property to characterise kind members, some of our intuitions concerning the nature of 'profound' classifications would be violated; intuitively, the charged bodies for instance do not constitute a natural kind; see Bird and Tobin (2008: section 1.1). That is, kind members should share several properties. That is why I shall often speak about patterns of properties (or of similarities) and not just about properties *simpliciter*.

the fundamental divisions of nature', etc.[2] Of course, these intuitions have played their role in the debates about what kinds are natural and what criteria and requirements for natural kind membership we should accept.

In medicine we also deal with classifications; diseases and diseased organisms are classified. In the present thesis, I focus on the life threatening conditions in somatic medicine and ask - could diseases be natural kinds? The abovementioned intuitions tell us 'no', by all means. Surely, classifications in physics or chemistry intuitively delineate the 'basic divisions of nature' and the fundamental particles or the chemical elements do seem to fit into a 'basic ontology'. Diseased organisms, in contrast, appear as second-hand relata of medical classifications, which, even though they might be non-arbitrary, do not seem to 'carve nature' deep enough, down to its 'joints'. In short, an 'ontological gap' seems to separate the natural kinds of the exact sciences and the disease kinds, as it were. Moreover, the fact that the very objectivity of medical classifications has been under dispute[3] - in line with more or less recent allegations that a value component is involved in the disease concept – seems to make the case for diseases as natural kinds even worse.

The main claim I will try to defend in the first part of this thesis is that notwithstanding these intuitions (and the problems concerning the exact disease/non-disease distinction), the ontological gap putatively separating kinds in the exact sciences from medical kinds is un-warranted. I will try to show that instead of an ontological gap, we only have a difference of degree. Hence, we can justifiably view diseases as natural kinds.

This conclusion will have a certain proviso, related to the question of whether membership of a kind is necessary for the identity of natural kind members. Only if their identity depends on the instantiation of kind properties, the caveat says, diseases could not be natural kinds. That is simply because the identity of organisms could not possibly depend on their diseased traits. On the other hand, although in principle the kind membership in physics or chemistry could influence the identity of the respective kind members, I will indicate a sum of epistemological reasons for why such a possibility is very implausible and difficult to assess.

---

[2] See for example Hirsch (1997: 52)
[3] See for instance Szasz (1987) Fulford (2001)

Methodologically, I will undertake my discussion by way of comparing a kind of disease (namely Graves' disease) with an exact science kind (namely Gold). I will show that no matter what putative criterion of natural kind membership we take into account, a warranted ontological gap does not show up between them. Three possible scenarios will appear in my comparison. For some criteria, both Gold and the Graves' disease kind will turn out to satisfy them. For other criteria, neither Gold nor the Graves' disease kind appears to be suitable. For a third group of criteria, only one of the two kinds seems to satisfy them but then we can seriously question the justification these criteria; the identity caveat mentioned above will play a crucial role in this questioning task.

The second part of the thesis continues the work of the first in the direction of scientific explanations. Obviously, the pragmatic reason why classifications abound in sciences is that they smooth out the way towards explaining the behaviour of classified particulars. In general, natural kinds are associated with causal explanations. That is, the behaviour of a particular is explained by pointing out that the properties characteristic of its natural kind cause it to behave as it does, in a *positive* way. In medicine, however, a large part of the explanations in use do not take an overt causal form but appeal to biological dys-functions (or failures to function). In other words, in such explanations the *absence* of certain functional effects is responsible for the state of disordered organisms. The issue I am concerned with in the second part of the thesis is how to reconcile this (dys-)function based register of medical explanations with the assumption that diseases are natural kinds and that the behaviour of ill organisms should be dealt with, on an explanatory level, by indicating the *positive* causes characteristic of (or identifiable through) their kind.

The suggestion I will advance is that instead of (or complementary to) identifying the dys-functioning of biological items, we should (also) seek to identify in pathological contexts their functioning, *per se*. Further, we can provide kind-specific explanations if we interpret the functioning as the manifesting of causal powers. Against the background of the Humean/ anti-Humean debate, I will show that Nancy Cartwright's capacities represent a plausible variant for the powers at work in 'pathological' functioning.[4]

---

[4] Cartwright (1999)

I should say that my enquiry into medical kinds and explanations is set out in a particular research context. Just as all the other special sciences, medicine has received in the last few decades a great deal of attention from philosophy. The philosophers discussing medical issues, however, have not primarily sought to explore the particular relevance for medicine of classical notions employed in the philosophy of science (notions like natural kinds, induction, laws, causation, explanation, etc.), but have rather concentrated on ethical aspects - for instance, the values that might be involved in identifying diseases. This focus of research has left little room for work on the metaphysics of medicine, as such. The latter has remained a relatively uncharted territory and the few authors who have ventured to discuss it were hampered by the absence of a background of systematic, concerted enquiry. Thus, for the most part, they could only import ready-made conclusions drawn elsewhere in the philosophy of science.

In this respect, the issue of function-based explanations is a good example. Theorists have appealed to functions as a possible means to draw the border between the pathological and the non-pathological, the claim being that one can identify diseased organisms via biological dysfunctions or failures to function of their biological items. Apart from this, however, the literature hardly offers any discussion of the explanatory relevance of biological functions in medicine. The related issue of natural kinds is also illustrative. There have been a number of papers and book chapters enquiring into the natural kinds of diseases, sometimes with interesting and provocative results.[5] Nonetheless, they have not really engaged with the current work on the metaphysics of natural kinds.

In a sense then, the central question of my thesis can be formulated as follows: if one sets aside the issue of the disease/non-disease distinction (and all the related discussions in which various ethical aspects are involved), what can be said, strictly vis-à-vis the metaphysics of medicine, about the kinds and explanations that the organisms suffering

[5] See Cooper (2005), Reznek (1987)

from life-threatening conditions are subject of?[6] Of course, in order to avoid the tempting range of ready-made conclusions, my answer will require a separate discussion of some general arguments from the philosophy of science about the nature of natural kinds. These arguments concern the problems of essences, hierarchies, substantive universals and processes. More deeply, they concern the sum of intuitions I have mentioned above, which influence kind theorists in backing up a particular set of criteria for natural kind membership.[7] My approach will be to offer two possible substantiations of the intuition(s) at hand and to question that particular set of requirements vis-à-vis their (made) explicit, intuitive grounds.

It is precisely that set of requirements that, it will turn out, poses the greatest problems in my methodological comparison. However, the problems of essences, hierarchies, substantive universals and processes, I will show, are quite innocuous when it comes to viewing diseases as natural kinds. If carving nature at its joints does *not* mean that the diachronic identity of natural kind members depends upon their membership, then it does not matter whether medical kinds satisfy or not that set of requirements. If carving nature at its joints *does* mean that the diachronic identity of natural kind members depends upon their membership, then an ontological gap could be in place between medical kinds and exact science kinds (as my caveat says, indeed). Nonetheless, we face huge epistemological problems when attempting to determine which properties have an identity-influence over the members of exact science kinds.

---

[6] I should underlie that I consider these ethical aspects extremely important and my decision to leave them aside is simply an issue of focusing on a particular path of enquiry. It is extremely important for instance, to come to terms with the problem of borderline cases in medicine – those conditions about which there is disagreement over their status as diseases. By focusing on life-threatening conditions, my enquiry does not demand (or attempt) a solution to the problem of how to draw the distinction between disease and non-disease in borderline cases (the interested reader should consult Margolis (1976) and Boorse (1977) for two representative defences of the value-ladeness and value-free positions, respectively). My views on this issue, vis-à-vis natural kinds, closely follow Cooper's approach from her (2005). Cooper maintains that a value component is involved in the disease concept but points out that this should not deter us from considering that at least some of the cases that medics treat, research, etc., qualify as natural kinds. A useful analogy drawn by Cooper is with the case of weeds and the weed concept – it might be that considering any plant a weed is a value judgment, but that being said, the plants put under this heading (e.g. daisies, buttercups) could be investigated as natural kinds; see Cooper (2005: 76). My account of diseases as natural kinds should be readily applicable to those conditions (among the borderline ones) that would conclusively be found to be diseases.

[7] In chapter 1, I will call these criteria the 'N&S requirement', 'non-overlapping requirement' and 'non-phase requirement'.

In outline, the structure of my thesis is as follows. In **chapter 1,** I present the various criteria and requirements for natural kind membership advanced in the literature and outline certain accounts of natural kinds adopted by particular philosophers that will play an important role in my arguments (Ellis, Lowe, Bird, Dupré, Boyd). I then set out the state of the art in the debates over the existence of natural kinds in medicine. In the final part of the chapter, I discuss the issue of nominalism vs. realism for natural kinds vis-à-vis the presented requirements for membership and then introduce the two substantiations of the carving nature at its joints intuition.

The problems of essences, substantive universals and processes are approached in **chapter 2**, where most of my discussion about the general metaphysics of natural kinds is located. I show that if carving nature at its joints does *not* mean that the diachronic identity of natural kind members depends upon their membership, then some of the more demanding criteria for natural kind membership that theorists have put forward in the literature cannot be justified.

In **chapter 3** I introduce the methodological comparison between gold and the Graves' Disease kind and take into account, in turn, all the criteria of membership presented in chapter 1. I then argue, in conjunction with the analysis undertaken in chapter 2, that there is no 'ontological gap' between kinds of diseased organisms and kinds in the exact science, maintaining however, the proviso that if the identity of natural kind members depends on their instantiating natural kind properties, then diseases could not possibly be natural kinds.

In **chapter 4** I turn to the other aspect of my enquiry and look at how we could explain the behaviour of diseased organisms, taken as members of natural kinds, from the point of view of biological functions. I argue that what we should expect from function-based explanations of diseased organisms, when viewed as members of natural kinds, is to illuminate positive causes of the symptoms and pathological behaviours they exhibit. I show that one can construe function-based explanations as fully-fledged explanations that refer to positive causes, if we interpret the functioning of biological items as the manifesting of causal powers. I also provide as a background for this discussion a presentation of the Humean/anti-Humean debate on causation and show next that Nancy

Cartwright's capacities represent a plausible variant for the powers that are at work in 'pathological' functioning.[8]

I argue that these capacities can be tracked down if healthy organisms are viewed as 'nomological machines', in the sense in which Cartwright understands this notion. I also suggest that capacities are necessary in order to vindicate general and, more importantly, singular causal claims involved in medical diagnosis and hence to found satisfactory functional explanations.

The thesis ends with some 'Final Remarks' in which I take stock of the arguments about medical kinds and explanations espoused in the previous chapters, situate tentatively the position of my thesis in the evolution that certain concepts have had in the history of philosophy and explain some of the argumentative strategies I have employed.

I should mention one final aspect in this Introduction. I will have to make a number of assumptions in my enquiry, having no space or scope to discuss their justification. I will specify each of these assumptions in the thesis, where their use is relevant for the discussion. For the purpose of clarity it is evidently useful to list them from the beginning as well. First, I assume non-reductionism of biological properties to physical and chemical properties. Second, I assume that the problem of (causal) induction is a metaphysical issue that has to do primarily with the existence or the non-existence of causal powers. Finally, I assume that metaphysical disputes can at most be clarified, but not solved by semantic and linguistic considerations.[9]

[8] Cartwright (1999)

[9] See especially Salmon (1982), Mumford (2005) and Lowe (2008) who discuss the issue of essentialism and the Kripke/Putnam semantic approach to it.

CHAPTER 1. NATURAL KINDS

INTRODUCTION

What are natural kinds? In scholastic philosophy, what we now call 'natural kinds' were supposed to be Ideas in God's mind, paradigmatic 'entities' that pre-existed and (thus) shaped the God's creation of the world of particulars.[10] In contemporary philosophy, the definition of natural kinds is less straightforward, if not slightly elusive. At base, natural kinds are meant to be metaphysical categories reflecting the fact that objects resemble one another in different degrees and some of these resemblances seem more 'profound' than others. The members of natural kinds, accordingly, should be individual objects that share (a) certain (pattern of)[11] 'important' properties, which we should consider in an 'ideal' classification. However, what are the 'profound' resemblances, what is an 'ideal' classification and what (patterns of) properties should count as 'important'? We have in the literature a great variety of answers to these questions, both on an explicit and intuitive level.

On an explicit level, we have a series of requirements for natural kind membership that aim to specify what nature the (patterns of) 'important' properties have. These requirements refer to inductions, laws, identity, universals, causation etc., and connect the natural kinds theme to various other controversial problems in metaphysics (such as the problems of essences, hierarchies, substantive universals and processes). On an intuitive level, we have a series of appealing renditions for what a 'perfect' classification should realize, e.g. 'carve nature at its joints', 'provide a basic ontology', 'divide the world in a non-arbitrary way', 'illuminate what things are', 'classify objects independently of our thoughts, values, conventions, language, etc.' 'delineate the fundamental divisions of

[10] See Funkenstein (1989)

[11] Again, as Bird and Tobin note, if there were a single property to characterise kind members, some of our intuitions concerning the nature of 'profound' classifications would be violated; intuitively, the charged bodies, for instance, do not constitute a natural kind; see Bird and Tobin (2008: section 1.1). That is, kind members should share several properties. That is why I shall often speak about patterns of properties (or of similarities) and not just about properties *simpliciter*.

nature', etc. [12]

Notably, there is a basic disagreement in the literature as to which criteria are 'correct' and the precise, intuitive content suggested by phrases like 'carving nature at its joints' is also unsettled (and often unquestioned). The two matters might well be connected. To disagree over the criteria of natural kind membership is in the end to disagree over natural kind Realism and the elusiveness of the 'carving nature at its joints' intuition might well reflect the tension faced by contemporary metaphysics when employing age-old conceptual schemes (in which 'natural kind' Realism was, as I said, a much more straightforward matter).[13]

My purpose in the present thesis is not to analyse this conceptual tension (if any) from the perspective of the history of philosophy and even less to try to settle what natural kinds are. However, I will have to undertake a discussion, with analytic, contemporary means, of the relationship between kind Realism and criteria of natural kind membership on the one hand, and the intuitive baggage of the contemporary natural kind discourse, on the other. That is because, as I have mentioned in the Introduction to this thesis, a crucial reason for dismissing diseases as natural kinds is that they seemingly could not form a 'basic ontology', 'show what things are', etc. and some requirements fulfilled by exact science kinds and failed by medical kinds have an intricate grounding in this intuitive luggage. I will undertake in chapter 2 the bulk of this discussion, which will revolve around the (possible) role played by natural kind membership for the diachronic identity of kind members. In the present chapter, I will introduce some key concepts and lay down both the background for the general, metaphysical discussion from chapter 2 and the more applied enquiry from chapter 3 in which the particular context of medical kind is investigated.

Thus, I will present in this chapter the various criteria and requirements for natural kind membership advanced in the literature (§1.1) and will outline certain accounts of natural

[12] See for example Hirsch (1997: 52)

[13] As it is well known, 'carving nature at its joints' is originally a platonic phrase (*Phaedrus*—though Socrates's words are somewhat different) that was taken up and exploited to the maximum by medieval philosophers. For some general analyses of the abovementioned kind of tension see Funkenstein (1989) who looks at the heritage of our notions of space, will and knowledge, Koyre (1957) who looks at the heritage of our cosmological concepts and Hacking (1984) who traces back the origin of the modern concept of probability.

kinds adopted by particular philosophers that will play an important role in my arguments (§1.2). Then I will set out the state of the art in the debates over the existence of natural kinds in medicine (§1.3) and will portray the dilemmas posed for natural kind Realism by the presented requirements for membership, as well as my own methodology vis-a-vis natural kind Realism (§1.4). In the final part of this chapter I will introduce what I see as two possible substantiations of the 'carving nature at its joints' intuition, in relation to the diachronic identity of natural kind members (§1.5).

§ 1.1 CRITERIA FOR NATURAL KIND MEMBERSHIP

All accounts of kinds claim that members of a kind are similar. However, questions remain with respect to what sort of *properties* kind members should share and of what types of similarity *patterns* they should exhibit. I will try in the present section to systematise a number of requirements advanced in the literature regarding the natural kind membership. The criteria at hand provide various answers to the questions above.

I – *natural properties requirement*

The first criterion tells us that properties delineating kinds should be natural. The reason why our classifications are arbitrary in some cases is simply that the similarities grounding these classifications amount to gerrymandered properties. The salient contrast at work here is between natural properties and so-called Cambridge properties – arbitrary, relational properties envisioned by manoeuvring predicates e.g. being 5 km away from the Eiffel tower. Evidently, the objects that are 5 km away from the Eiffel tower are less likely to be candidates for natural kind membership than, say, the objects made out of aluminium. Philosophers who deny this requirement affirm in turn that any classification is just as good as any other. Goodman makes an adequate example here since his paradox of induction unveiled using grue-predicates (a kind of Cambridge

properties) questions the status of natural properties.[14]

Sometimes theorist contrast natural properties with relational properties in general.[15] It is debatable though whether all relational properties should be put into the same category as the Cambridge ones. There are salient relational properties, most remarkably the ones found in biology (e.g., lineage), that are far from being gerrymandered and arguably should not be ruled out in the first instance as candidates for being instantiated by natural kind members.[16]

Insofar as properties are natural, one could rule out a source of arbitrariness for our classifications. Indeed, this requirement concerning natural properties is often formulated by saying that our classifications should be *non-arbitrary* in the sense that they should not depend on our language, values, conventions, thoughts, etc.[17] Sometimes, non-arbitrariness is also cited in connection with the rest of the membership criteria, to be listed below.

II - *causal induction requirement*

Besides being characterised by natural properties, what is also required from kinds is to participate in consequential, causal inductions.[18] For instance, it might well be that the kind water – characterised *as* liquid, transparent, dissolving sugar, etc. - is delineated by natural properties. Nonetheless, *this* kind does not participate in causal inductions in the same sense as does the kind delineated by more 'profound' properties (like covalent and hydrogen bonding between atoms of oxygen and hydrogen, which dissociate under this or

---

[14] See Goodman (1946) and Goodman (1965) in which not only natural properties but also all our metaphysical categories are relativized upon the holding of certain conceptual schemes.

[15] Ellis (2001) and Wilkerson (1995) among others, require the properties in question to be *intrinsic*. Ellis notes and discusses the problems with the exact distinction between intrinsic and relational (Ellis, 2001: 26-30); see also Bigelow (1999: 52, 53)

[16] See Dupré (1993) who advocates the use of relational properties (in biology)

[17] See Hirsch (1997: 53-60), Bird and Tobin's discussion of *naturalism* in their (2008) and also Hacking's review of Mill's kinds in Hacking (2006)

[18] A criterion put forward by, for instance, Wilkerson (1995: 32), Ellis (2001: 4-7 *passim*), Mackery (2005: 446), and Boyd (1995: 368). As I have stated in the Introduction, I assume that the problem of (causal) induction is a metaphysical one, which has to do primarily with the existence or the non-existence of causal powers.

that ionic product, with the corresponding V shaped structure having a high density electron distribution around the oxygen atoms, etc).[19] Of course, both sets of properties uniquely pick out the kind water (in the actual world at least). Nonetheless, the latter sort of properties, unlike the former, is causally efficient - they act as causes for other 'superficial' properties we seek to explain. The reason that inductive conditionals of the form "if something is a sample of water, then it will, or it probably will, boil at 100°C, under normal atmospheric pressure" are successful is because the water kind being is characterised precisely by such explanatory powerful properties: the sort of properties that science makes reference to in scientific explanations.

Hence, a kind supports consequential causal inductions when the properties that characterise it are 'profound' from a causal point of view - properties attended by causal powers, with roles in scientific explanation, prediction, etc. For simplicity, I will henceforth call these properties *determining* properties.[20] Evidently, most often these determining properties that science makes use of are microstructural properties and the properties they explain are observable ones. However, it is worth stressing that, insofar as the criterion of (causal) induction is concerned, the distinction between the 'profound', determining properties and the 'superficial' properties has strictly causal grounds and is thus different from the distinction between microstructural and observable properties. That is because, for one thing, the 'superficial' properties we should explain might be microstructural properties as well, and for another observable properties and behaviours may have causal powers as well.

III - *law involvement requirement*

A bolder development of the previous criterion says that natural kinds should be involved

[19] See Chaplin (2008)

[20] I follow here Cooper's (2005) terminology. Note that certain authors (especially those following the Lockean discussion of 'essences') call these explanatory powerful properties as 'essential'; see for instance Harre (2005: 10, 11). I shall return to different senses of the essential shortly.

in (causal) laws of nature.[21] It might be for instance that electrons are involved in induction inferences of the type 'if all electrons observed so far exhibited this or that behaviour towards protons, then the next observed electron will, or will probably exhibit the same behaviour', or 'if something is an electron, then it will attract protons with this force'. However, one could point out that what is at work here is a *law* in which the kind of electrons is involved, because *charge* is one of its characterising properties.

In other words, this criterion says that the properties shared by kind members should be not only natural and explanatory powerful but also 'nomic' – the sort of properties that would be found in the antecedents of law statements, were such statements formulated by appealing to property-predicates.

Note that this is a correct rendition of the requirement that kinds should be involved in laws only if kinds do not *found* laws, i.e. only if the casual mentioning of kinds in law-like sentences is not interpreted metaphysically in the sense that kinds are construed as truth-makers of law-statements.[22] The mainstream view, however, is that the relata of laws are properties. Note also that this criterion is different from criterion II, insofar as kinds could participate in causal inductions without being involved in laws, and there are serious reasons to think this is the case. For instance, special sciences represent for many scientific law theorists domains in which we do not have laws, at least in the sense in which canonically, laws characterize the exact sciences.[23] Yet, insofar as one does not advocate reductionism (for instance of the biological or psychological to the physical level), special sciences cannot be understood as lacking causation, i.e. as being characterized by arbitrary, accidental processes. Hence, in such domains it should be meaningful to discuss about explanatory powerful properties and causal inductions, even if the existence of nomic properties, in the full-blooded sense, is an undecided issue.[24] I will return to the topic of kinds and laws at various other points in this chapter, especially

---

[21] Criterion put forth by Bird (1998), Collier (1996), Ellis (2001) among others. One should distinguish between causal and non-causal laws – what Hempel for example called 'laws of succession' and 'laws of coexistence' (Hempel, 1966). At work in the above criterion, which strengthens the causal induction criterion, are the *causal* laws. The co-instantiation laws are involved in what I shall call the N&S requirement (criterion VI_b – which I will describe below).

[22] See Lowe (1989) who adopts this view of natural laws.

[23] See Ellis (2001) and Carnap (1995).

[24] See for instance Boyd (1991), who discusses persuasively about causal induction and causation in biology while accepting that the type of laws found in physics have no correspondent in the biological realm.

in §1.2, and in chapter 3.

IV - *identity requirement*

A final condition set upon the properties characterising a natural kind is for them to be involved in the identity of the kind members, in the sense that, should a kind member lose its membership, it would cease to exist. The underlying view here is that the (numerical) identity of a particular depends not only on spatio-temporally continuity or a certain causal history, but also on its continuing to instantiate the natural kind properties. For example, in spite of the fact that the microstructural properties of lead are natural, explanatory powerful, nomic, etc., in order for lead to be a natural kind it should be the case that, were a particular sample of lead transmuted into gold, the end-result would be numerically different from the sample before the transubstantiation.[25]

Now, the previous four criteria spell out what sort of *properties* should be instantiated by kind members. We have, however, a further set of requirements concerned with the *co*-instantiation of properties by kind members and the nature of the *patterns* of properties they share. These additional criteria complement the first four in the sense that the properties considered as parts of the respective patterns should be at least explanatory powerful, determining properties.

V - *co-instantiation induction requirement*

One such criterion is that natural kinds should be involved in co-instantiation inductions, which we should distinguish from causal inductions (just as we should distinguish causal laws from co-instantiation laws). For example, if something is a metal, then it will (or it

[25] The proponents of this criterion are Brody (1974), Denkel (1996), Wiggins (1980), Elder (2005, 2007) These identity-bearing properties are also at times called *essential*; evidently, in a different sense than the one associated with the explanatory powerful, determining properties. Parenthetically, if we reckon with the distinction between countable and non-countable (or mass) kinds, in the case of the latter this criterion entails that their members' diachronic identity has changed after losing their kind membership. See Lowe (2001: 73, 74) and Lowe (1998: 199-201) for his related distinction between individuals and *pseudo*-individuals.

will probably) have the valence shell electrons delocalised, and have available a far larger number of delocalized energy states than of delocalized electrons, and possess a lattice of positive ions, etc.

I have said that the properties involved in this additional set of criteria concerning the nature of the *patterns* of similarities exhibited by kind members (and accordingly, involved in the co-instantiation induction criterion as well) should be at least determining properties. This observation is particularly relevant here because it helps us to differentiate between the co-instantiation induction involved in the present criterion and the causal induction involved in criterion II.

Most if not all co-instantiation inferences that can be framed using (causally) 'superficial' properties – e.g. if something is a duck, then it has webbed feet, a long neck, serrated lamellae of the bill, etc. – are masked causal inductions that appeal to the presence of certain determining properties (in this case, the genetic properties of ducks). Such masked causal inductions are in fact dependent upon prior co-instantiation *inferences* concerning determining properties, the latter inferences being much more genuine candidates for co-instantiation *inductions*.[26]

VI_a - *cluster requirement*

Relative to requirement V, we have two separate criteria concerned with the limits of the properties' patterns shared by kind members (or, differently put, concerned with the degree of similarity required for kind membership).

One criterion says that we should identify the limits of these patterns as the regions of high density in the logical space of properties, corresponding to individuals that share a certain *cluster* of properties. In order for an individual to qualify as a natural kind member, it should instantiate (all or) some of the properties of the cluster. Different parts of the cluster could be instantiated and no part should be necessary and sufficient for

[26] Of course, in theory, a causal story could be brought into discussion for the case of co-instantiated determining properties as well (and not only for the co-instantiation of 'superficial' properties), as we shall see in the following.

kind membership. The rationale that attends this requirement is that, due to the causal structure of the world, properties tend to agglutinate; by citing either (external) mechanisms or (external) laws, we could vindicate the co-instantiation inductions.[27]

Boyd (1991), who focuses mostly on biology, advocates *mechanisms* as the means through which the 'causal structure of the world' produces the clustering of properties. Bird (2007) focuses on the exact sciences and proposes *laws* instead. On Bird's proposal, the kinds in physics and chemistry would end up being non-fuzzy, but simply in the sense that laws would produce a perfect agglutination of properties, resulting in 'perfect', well-cut patterns.

It is worth noting that this criterion allows kinds that might be fuzzy or overlap each other, in the sense that natural kind members might instantiate different properties and two different natural kinds might share members, even if they were not in a species-genus hierarchy.[28]

VI_b – *N&S requirement*

The other criterion (standing somehow in opposition with VIa) concerned with the degree of similarity between kind members, says that every natural kind should have a set of properties that are necessary and jointly sufficient for kind membership.[29] Notably, theorists avoid sometimes the 'necessary and sufficient' idiom when discussing this requirement. The reason is that the necessary and sufficient conditions of membership have long been associated with a set of 'superficial', 'observable' properties considered

---

[27] See Boyd (1991: 129), Bird (2007: 208-211) Kornblith (1995: 35, 43). Dupré, in his (1993), sticks to the cluster requirement but does not mention the underlying, causal justification for it. This diminishes much of the coherence of Dupré's argumentation, or at least of its explanatory scope. On a different note, I have called mechanisms and laws (ensuing the co-instantiation unity of the patterns of properties) 'external' from a *metaphysical* point of view, in contrast with the 'internal' sense in which substantive universals (or 'essences' *simpliciter*) are supposed to ensure such co-instantiation (see criterion VIb above). In biological cases, from a *physical* point of view, mechanisms could be both internal and external (i.e. within a certain organism or outside of it) but they would still be metaphysically external to the pattern of properties they agglutinate (because of latter as the truth-makers of the former. Otherwise, I believe, laws (irrespective of whether they are causal or co-instantiation laws) are external to kinds, *pace* Tobin (2008: 9)

[28] See Dupré (1993: 18 *et passim* efficient causation). Similarly, laws could be metaphysically 'internal' to kinds only if we construe the.)

[29] See Wilkerson (1995: 109) and Harre (2005: 9,10)

potentially useful for *recognising* natural kinds but not fundamental for their very nature; the preferred formulation of this criterion, a negative one, is that natural kinds should not be fuzzy. There should not be any gradual transition between one kind and another, certain authors claim, for instance - where this formulation, we could easily infer, simply precludes the interweaving of clusters that might show up when no necessary and sufficient conditions are in place.[30]

This criterion (which, nevertheless, I will call 'N&S requirement' henceforth) has an evident connection with the co-instantiation induction criterion, even though sometimes authors may advance it for independent reasons.[31] The necessary and sufficient conditions are supposed to delineate 'essences' which, informally put, should constitute an 'internal' glue and thus account 'internally' for the co-instantiation of properties by kind members. As to what metaphysical category the essences should belong to, one direct answer is that we have essences *simpliciter,*[32] or, in a more precise interpretation, that we have substantive universals which are instantiated by kind members.[33]

---

[30] The formulation belongs to Ellis, who advocates accordingly the 'categorical distinctness' that should be in place between natural kinds. At other times, Ellis speaks about the exact resemblance, or the 'identity' between different natural kind structures (*Cf.* Ellis, 2006: 67-68). In this context, Ellis says that the properties of the clear-cut patterns are necessary and sufficient for the *identity of the kind* (Ellis, 2001: 19, 52, 75, 246). This is a third sense in which the kind properties are sometimes called *essential* - besides the sense in which properties should be explanatory powerful and the one in which they should have an identity bearing on kind *members*.

[31] Ellis thinks that if there were a gradual transition between kinds, the limits between them would have to depend on us, on our decision to draw a dividing line. Kinds would then turn out to be arbitrary. (*Cf.* Ellis 2001: 19)

[32] Many authors mention Lockean Real Essence at this point. Parenthetically, the invoking of Lockean real essence constitutes a great source of equivocation between three senses of 'essential' properties, to wit, as explanatory powerful, as essential-for-kinds (or necessary and sufficient for the identity of kinds) and as essential-for-kind members (necessary for the identity of kind members; see Wilkerson (1995: 30,55) for instance, or Brigandt (2009), Harre (2005). Ellis notes the distinction between the two senses by observing that kinds with Lockean essences might be fuzzy (Ellis, 2001: 31) but in his arguments as such the distinction is often overlooked (see Ellis, 2001: 19-38). I shall return to this aspect in chapter 2. Note that there is also a fourth important sense of the essential, recently (re-)introduced in the natural kind discussion, namely that essential properties could be necessary for kind membership (Bird, 2009). I will not discuss it in the present thesis though, because, *inter alia*, this fourth sense of the essential is consistent with accepting requirement VIa and denying requirement VIb.

[33] Both Ellis (2001) and Lowe (1989), (2001) view kinds as substantive universals. There is an important difference between their views: for Ellis the *members* of natural kinds should instantiate all the properties that are necessary and sufficient for membership (kind identity, in his idiom). For Lowe, these properties (at the level of their corresponding non-substantive universals), *characterise* the natural kind substantive universals, but do not have to be instantiated by all kind members. *Cf.* Lowe (2006: 26)

VII - *non-overlapping requirement*

Closely connected with the above VI_b criterion is the non-overlapping requirement, which states that natural kinds may not overlap each other. That is, if two natural kinds share members then the kinds are either identical or are part of a species-genus hierarchy.[34] I should stress that in principle, we could have non-overlapping kinds that present fuzziness and non-fuzzy kinds that are overlapping; that is, we could have kinds that respect criterion VII and violate criterion VIb and vice-versa. However, kind theorists discuss so closely the two criteria[35] that I will not insist in general upon exceptional cases. In §2.1.2, however, when discussing a certain approach to the 'essences' of natural kinds, I shall take into account the fact that there is no relation of logical entailment between criterion VIb and VII.

VIII - *non-phase requirement*

Finally, what is required from the patterns characterising kinds is that their instantiation should not constitute a phase in a process that is parasitic upon a (more stable) fundamental pattern. In the following, I will call this the non-phase requirement.[36] The requirement in question rules out for instance that tadpoles (in contrast with frogs) or ice (in contrast with water) could form a natural kind and it is intricately involved in the debates over the status of natural kinds in biology, from the point of view of evolution. It also has a direct relevance for our topic insofar as diseases could be construed as processes (that come to be) suffered by 'healthy' organisms. I will address this point in chapters 2 and 3.

To sum up, I have presented in the present section various membership criteria advanced in the literature. I have split them into criteria concerning the nature of properties (the first four of them) and, complementary, criteria concerning the nature of the patterns of properties natural kind members should possess. Figure 1_a lists all these requirements.

---

[34] See Ellis (2001), Thomson (1969), Wilkerson (1995), Hacking (1993) who endorse this requirement.
[35] The most salient example being Ellis (2001); see Khalidi (1998: 40,41,50) for a (critical) discussion
[36] See Lowe (2001: 174-178), Wiggins (2001: 33 *et passim*). Often, this criterion is not stated as such, but is alluded to in the claim that biology could not have natural kinds, because of the evolution process.

They are also tentatively organised in 1_b according to their increasing degree of complexity and/or inter-dependence – namely, top to bottom in the order of complexity for the lower part of the figure (concerned with the *property* criteria) and bottom to top for the upper part (concerned with the *pattern* criteria).

***FIGURE* 1 CRITERIA FOR NATURAL KIND MEMBERSHIP, SPECIFYING WHAT NATURE THE PROPERTIES AND THE PATTERNS THEY FORM SHOULD HAVE (the dotted lines indicate the increasingly complex demands placed by these criteria over the nature of the natural kind properties and the nature of the patterns they form)**

a. Criteria for natural kind membership

I. Instantiation of natural properties (the non-arbitrariness criterion)

II.* Instantiation of explanatory powerful, determining properties (causal induction criterion)

III. Instantiation of nomic properties (law involvement criterion)

IV.* Instantiation of identity-bearing properties for the kind members (identity criterion)

V. Co-instantiation induction requirement

VIa. Cluster requirement

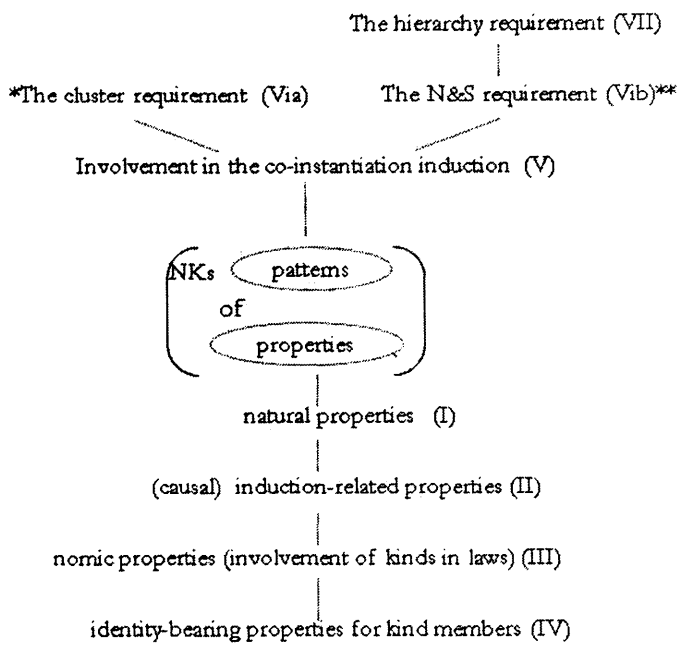
Vib.* N&S requirement

VII. Non-overlapping requirement

VIII. Non-phase requirement

* The marked criteria are associated with a construal of properties as *essential*, in different senses though – explanatory essential (criterion II), identity bearing for kind members (criterion IV), identity bearing for kinds (criterion VIb)

b. Tentative organisation of the natural kind criteria



*What ensures that properties cluster are either laws or mechanisms (which are metaphysically 'external' to the kinds

** The properties considered as necessary and sufficient are described as delineating 'essences' (which ensure 'internally' that a stable co-instantiation is in place). The category of essences is either unspecified (we have essences *simpliciter*) or is identified with that of substantive universals.

I did not mean to provide in figure 1_b a definitive organisation of kind membership criteria, such that any other possibility of ontological dependencies between them exists. For instance, even if the induction-related properties need to be at least natural and the

nomic properties need to be at least induction-related, one could see the N&S requirement not (only) as going a step forward from the criterion of involvement in co-instantiation induction, as my 1_b suggests, but (also) as a condition for our classifications being non-arbitrary. However, even though the interdependencies between these criteria could be differently put, I think that the organisation in 1_b reveals an important sense that makes all these requirements part of a unitary picture.

In figure 2_a below, I have introduced a simple notation for the content of the membership criteria; I will frequently appeal to it in the chapters 2 and 3. In figure 2_b the difference between the criterion VI_a and VI_b is drawn *vis-à-vis* the status of co-instantiated properties.

***FIGURE* 2 The criteria of membership in a natural kind NK for an individual *a* instantiation a pattern $[p_1, p_2, \ldots\ldots p_n]$ of 'profound' properties and a $[q_1, q_2, \ldots\ldots q_n]$ pattern of 'superficial' properties**

2a

I. $[p_1, p_2, \ldots\ldots p_n]$ should be natural properties

II. $[p_1, p_2, \ldots\ldots p_n]$ should be responsible for $[q_1, q_2, \ldots\ldots q_n]$ (as well as for the causal interactions with other individuals).

Were *a* to be NK, a would, or probably would, instantiate $[q_1, q_2, \ldots\ldots q_n]$

III. $[p_1, p_2, \ldots\ldots p_n]$ should be involved in laws (either governing/describing the causal interactions whose effects are $[q_1, q_2, \ldots\ldots q_n]$ or the interactions with other individuals

IV. $[p_1, p_2, \ldots\ldots p_n]$ should be involved in the (absolute) identity of *a*. Were *a* to stop instantiating any of the $[p_1, p_2, \ldots\ldots p_n]$ properties, it would not be the same numerically identical individual (or it would not have the same diachronic identity conditions, if *a* was a sample of a non-countable kind)

V. The co-instantiation induction criterion – were *a* to be NK, it would instantiate $[p_1, p_2, \ldots\ldots p_n]$

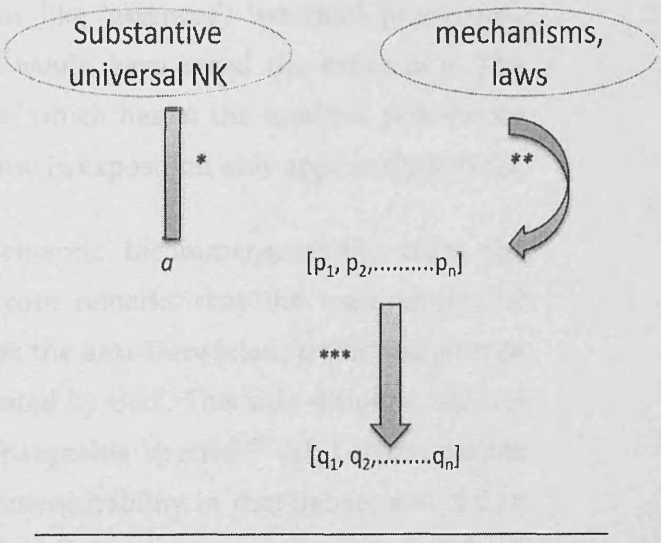
VIa. $[p_1, p_2, \ldots\ldots p_n]$ form a cluster due to the causal structure of the world; neither all of $[p_1, p_2, \ldots\ldots p_n]$ nor any part of them are both necessary and sufficient for membership

Vib. $[p_1, p_2, \ldots\ldots p_n]$ are necessary and sufficient for membership

VII Was *a* to be a member of a different natural kind NK', then either NK' would be identical with NK or the two kinds would be in a species-genus hierarchy

VIII. The instantiation of the pattern $[p_1, p_2, \ldots\ldots p_n]$ should not represent a phase in a process whose end result is another, non-phase pattern.

2b



* Between the individual *a* and the substantive universal NK there is a relation of instantiation, due to which a possesses the pattern $[p_1, p_2, \ldots\ldots p_n]$

**Alternatively, the possession by a of $[p_1, p_2, \ldots\ldots p_n]$ might be viewed as the result of the (efficient) causal structure of the world (mechanisms, causal laws)

*** Between the 'profound' and the 'superficial' properties there is a causally efficient relation. The arrow should indicate both this relation between the two patterns instantiated by *a* as well as the causal interactions with other individuals in which *a* engages. For these interactions, a subset of $[q_1, q_2, \ldots\ldots q_n]$ could be taken to designate 'observable' behaviours or dispositions, irrespective of whether the latter are considered intrinsic or relational properties.

Admittedly, the causal relations involving mechanisms/laws and determining properties, just as the causal relations holding between the determining and superficially properties are very much simplified in *2b* - see for instance Boyd's rendition of the role of mechanisms, to be presented in the following section. However, the respective figure is heuristic for the distinction between the causal and the co-instantiation induction, which will play a central role in chapter 3.

The next section will present a few of the most important accounts of natural kinds,

which employ some or all of these requirements and which will have a bearing on my discussion of the medical kinds in the following chapters.

I should finally add a point of clarification to this section. In presenting these requirements for natural kind membership, I have tried to avoid as much as possible the use of semantically loaded terms and locutions like 'essences', 'essential properties', 'species', 'being what one is', etc., even if they would have eased the exposition. The semantic load of the term 'essence', for instance, which has in the modern, post-Locke context at least four different senses,[37] makes its use in exposition only apparently helpful.

In his discussion of the case of apparent semantic incommensurability from the nineteenth century debate on Darwinism, Laporte remarks that the main source of strained communication was simply the fact that the anti-Darwinian, traditionalist side associated the concept of species with 'being created by God'. This side simply could not accept the rival theory couched in the terms of changeable 'species'.[38] Yet, Laporte doubts that there was a case of genuine semantic incommensurability in that debate and thinks that a communication could have been established if simply the term species had been dropped or if the parties had agreed to distinguish between the different parts of the manifold semantics each of them associated to the respective term.

To be sure, there is no case of semantic incommensurability in the natural kinds discussion, but nevertheless, I think that the scope of arguments would greatly benefit from an economical use of semantically loaded terms (as well as from a precise definition of their senses, whenever they are used). It is for related reasons that I chose to assign numbers to each criterion of membership earlier in this chapter, named (a bit ponderously) the similarities between kind members 'patterns of properties', and so forth. These choices might make the exposition slightly more difficult in the following chapters, but in return, I hope that the arguments will gain in clarity.

This is not to say that terms like 'species', 'essences' and phrases like 'carving nature at its

---

[37]And a few others, I suppose. For instance Murphy (2006) discusses the simple essences and refined essences; Brigandt (2009) speaks about origin essences, qualitative essences, epistemologically fundamental essences and essences as 'the feature determining an entity's identity' where the entities in question are species.

[38] *Cf.* Laporte (2004: 122)

joints' should not be used. Nevertheless, their senses should be clearly distinguished. A great deal of the discussion in chapter 2 of criterion VIb (N&S requirement) for instance, will revolve around distinguishing different senses of the essential and around disentangling these senses, as employed in the works of certain kind theorists. The case of 'carving nature at its joints', as we shall see in § 1.5, is another one in which what we mean should be plainly delineated.

§1.2 MAIN ACCOUNTS OF NATURAL KINDS

With the possible criteria for kind membership discussed, we now have a framework for considering the accounts of some key theorists. Their views are very different and diverse, as it can be seen preliminarily in *Fig.* 3 below, in which their allegiance to some specific set of criteria is represented. Their approaches thus cover almost the whole spectrum of views advanced in favour of one particular account or another.

FIGURE 3 KEY ACCOUNTS OF NATURAL KINDS

Criteria	I Natural property criterion	II Causal induction criterion	III Law involvement criterion	IV Identity criterion	V Co-instantiation induction criterion	VI_a Cluster criterion	VII_b N&S criterion	VII Non-overlapping criterion	VIII Non-phase criterion
B. Ellis	√	√	√		√		√	√	
J. Lowe	√	√	√		√				√
A. Bird	√	√			√	√	√		
R. Boyd	√	√			√	√	√		
J. Dupré	√		√			√			

Brian Ellis takes the notion of natural kind as a cornerstone of his argumentation designed to refute Humeanism and show that laws of nature are metaphysically necessary. Ellis argues that the *properties* characterising natural kinds should be natural, explanatory powerful (determining) and should participate in the (causal) laws of nature, rejecting however the condition that the identity of kind members depends on the kind properties they instantiate.[39] In other words, Ellis adopts criteria I-III described in the previous section and rejects criterion IV.[40]

Natural properties are contrasted with Cambridge properties and the former are defined as being intrinsic; this is one of the reasons why Ellis rejects the view that natural kinds exist in the biological realm. The explanatory powerful properties are mostly introduced by contrasting the microstructural properties of chemical elements and compounds with their observable features. The former are to be preferred to the latter, says Ellis, not only because the former are intrinsic whereas the latter are relational (the most often cited example such observable properties that are supposedly relational being colours), but also because the former are attended by causal powers, etc. and produce the latter.[41]

Ellis sometimes calls the 'properties attended by powers' (in the terminology I have used so far) simply 'powers' and divides the microstructural properties into two sorts: powers and so called structural properties, namely the properties underlying similarities in shape, spatio-temporal arrangement of various components, etc. The contrast between powers and structural properties is set out by using a certain construal of the famous dispositional/categorical distinction – powers should confer to their bearers the possibility of engaging in certain causal interactions whereas structural properties, as

---

[39] *Cf.* Ellis (1999: 67,68) I shall restrict my presentation of Ellis's views to the theme of natural kinds of objects, ignoring for the most part the other veins of his (extremely rich and impressive in its scope) metaphysical picture.

[40] Ellis specifies that natural kinds should be inductively rich but, as most authors, does not make the distinction between co-instantiation and causal inductions. His examples of laws include, however, co-instantiation laws and laws in turn are said to justify inductions.

[41] See Ellis (2001: 31). The background of the discussion is Locke's distinction between real and nominal essences.

categorical, should be neutral to such interactions.[42]

Kinds are involved in causal laws precisely because laws describe the manifestation of these powers (intrinsic properties possessed by kind members). These 'nomic' properties do not have to be involved, however, in the identity of kind members and therefore it is possible for objects to swap kinds while retaining their identity. One cannot see any distinction, says Ellis, between an atom losing an electron and its losing a proton, insofar as the implication of these changes on the identity of the individual atom in question is concerned.

As regards the criteria concerning the properties' patterns, Ellis adheres both to the N&S requirement and to the non-overlapping requirement (that is, to criteria VIb and VII). Natural kinds should not be fuzzy and there should be a genus-species hierarchy for any overlapping kinds. One reason for this is to avoid arbitrariness in classification. If kinds were fuzzy and hence there were transitions between them, then the line between one kind and the other would depend on us. If on the other hand kinds overlapped without being in a hierarchy, allotting an individual to one kind or another would hinge on our decision as well and would not reflect the 'divisions of nature'. These requirements also have a more elaborate metaphysical underpinning. In order to explain the 'identical' structures exhibited by natural kind members and the exact similarities between the powers they possess (i.e. in order to metaphysically justify the N&S requirement), Ellis introduces a class of substantive universals whose tropes should be instantiated by kind members. The powers that the kind members possess are also taken to represent tropes of universals, of a non-substantial type and to form 'essences'.[43]

Jonathan Lowe takes natural kinds as a crucial part of his four-category ontology that should reconcile diverse metaphysical positions about the nature of individual substances

---

[42] As Ellis frequently states – due to their powers, objects are disposed to behave in certain ways; see Ellis (2008: 140) Note that the categorical/dispositional distinction does not have a direct connection with our discussion. This distinction is mentioned in this section only insofar as it underlines the different approaches that the two most important proponents of kinds as substantive universals (Lowe and Ellis) have with respect to the way in which kind members come to instantiate (patterns of) properties.

[43] Ellis (2001:70, 73-75, 92) For the justification of the non-overlapping requirement, see Ellis (2001: 67,68), where the species-genus relation between substantive universals is introduced. I will discuss in more detail in chapter 2 how precisely the substantive and non-substantive universals should metaphysically justify the N&S requirement.

and their properties, following Aristotle's footsteps. Just like Ellis, Lowe construes kinds as substantive universals, and the patterns of properties possessed by kind members are similarly taken to correspond to non-substantive universals. There are important differences between their approaches though. Lowe introduces a relation of *characterisation* between the non-substantive universals and the substantive ones. This relation should explain the phenomenon of co-instantiation, in that the possession of certain properties (tropes of non-substantive universals) by individuals that are kind members (i.e. that are bearers of substantive universals) should reflect, at the level of particulars, the fact that the non-substantive universals *characterise* substantive universals.

Lowe accepts criteria I and II by stating that the properties possessed by kind members should be natural and explanatory powerful but does not employ Ellis's dichotomy between structural (geometrical) and non-structural properties. Lowe uses substantive universals to work out a different distinction between the categorical and the dispositional and thus to define the powers possessed by kind members. Lowe states that, from the point of view of universals, the predication of properties to kind members is dispositional, and indicates to the fact that non-substantive universals characterise substantive universals. Insofar as we strictly keep up with the level of particulars, the predication of properties to kind members is categorical (or occurent), and indicates that a trope is instantiated by a kind member.

In other words, in the occurent (or categorical) register kind members *actually* instantiate certain properties. In the dispositional register, kind members *can* instantiate certain properties. The difference between the dispositional and the categorical register is not (just) a difference in the mode of predication. It marks the metaphysical difference between a non-substantive universal that is in a certain relation with the kind universal and the trope of such a non-substantive universal instantiated by a kind member.[44]

Whereas for Ellis, powers are those properties whose instantiation *can* produce the subsequent instantiation of another property (the manifestation) in the same or in

[44] *Cf.* Lowe (2006: 17)

another individual,[45] for Lowe, powers are those properties that *can* be instantiated by kind members in virtue of their membership. These could be properties *simpliciter* (and Lowe cites as an example the case of the properties instantiated by the members of the kind electron) or properties involved in causal interactions (and Lowe's examples include the dissolving of salt into water, the rotation of planets into circular orbits).[46]

This construal of the dispositional/categorical distinction has also a bearing on the way in which Lowe accepts criterion III referring to the involvement of kinds in laws. For Lowe, laws simply describe the relation of characterisation from the level of universals. Hence, even if the properties that kind members possess are in a sense 'nomic', on a more fundamental level kinds are involved in laws *directly* (and that is why the latter are expressed in second order quantified statements). Finally, the special approach to the categorical/dispositional distinction motivates Lowe's rejection of the requirement VIb concerning the patterns of properties instantiated by kind members. These patterns from the level of particulars do not have to be clear-cut, simply because the respective *characterisation* relation from the level of the universals just entails that kind members (instantiating substantive universals) *can* instantiate tropes corresponding to certain non-substantive universals. The manifestation of laws at the level of particulars admits exceptions and this is in fact considered an explanatory advantage. Lowe's expository example is that of the law 'ravens are black' which, he argues, is not disconfirmed by the

[45]...and whose instantiation *actually* produces, 'in such and such circumstances E_{ij}', the instantiation of the manifestation/effect properties. See Ellis's preferred notation of the consequent of causal laws in his (1999:68-69)

[46] The sort of properties instantiated by kind members in the case of causal interactions (the properties that are commonly regarded as manifestations of dispositions, e.g. being dissolved, or being stretched) could be taken as intrinsic properties *simpliciter*; *Cf.* Lowe (2006: 40) where this reading seems to be allowed. These properties could also be taken as relational properties; they underlie processes and exemplify certain relations between substantive universals (*Cf.* Lowe 2006: 17, 161-163). For instance, the power to dissolve in water possessed by the members of the natural kind salt should be due to a relation between the water kind and the salt kind. This relation gets exemplified at the level of particulars when a particular sample of salt dissolves into a particular sample of water, so that the occurent predications involving the 'is dissolving' predicate become appropriate. Note that at this point Ellis introduces processes-universals that are connected with his power universals and *not* with his substantive universals; see Ellis (2001:4, 19-22) and also Clapp (2002: 592)

presence of albino ravens.[47]

Lowe rejects requirement IV as well - kind membership is not taken to have any bearing on the identity of kind members. Individuals could swap kinds (this might be a physical possibility or just a metaphysical possibility) while remaining the same individuals. However, the numerical identity of individuals is dependent on them remaining within the same (what Lowe calls) ontological *categories*. Ontological categories are divisions of individuals, above the level of kinds, which make it metaphysically impossible for an organism (e.g. a horse) to turn into an artefact (e.g. a statue) and for an immaterial, inorganic object (e.g. a bronze piece) to transform into a living thing.

These three categories of individuals (the categories of living organisms, material objects and artefacts) are more general than kinds and yet they have a direct connection with kind membership because they are used by Lowe to work out criterion VIII (regarding the phase/non-phase distinction).[48] For each category, Lowe circumscribes a certain set of processes that mark the transformation of a phase kind into a non-phase kind, or on the contrary, that qualify a transformation undertaken by a natural kind as just a phase one.[49]

Alexander Bird invokes natural kinds, like Ellis, in the context of an anti-Humean argumentation (which should prove that the laws of nature are metaphysically necessary). Bird argues (*contra* Ellis and Lowe) that substantive universals are not requisite for an account of scientific laws. Causal laws can be founded on the ontology of non-substantive universals only, according to which the causal interactions described/governed by laws reflect relations between non-substantive universals that are ingrained in universals'

[47]For such cases of exception-ridden generalisations, *ceteris paribus* clauses, attached to first order statements (and said to refer to *cp* laws) are often invoked in the literature. However, Lowe argues that without introducing substantive universals, such *cp* clauses cannot escape the charge of triviality (*Cf.* Lowe 1989: 153-154). Lowe also extends his critique of *cp* conditions to the first order quantified statements meant to indicate counterfactually in which conditions powers (dispositions) produce (are followed by) manifestations (*Cf.* Lowe 2006: 13-17)

[48] Lowe is the only author from the five presented in this section who discusses explicitly criterion VIII regarding the phase/non-phase distinction

[49] See Lowe, 1998: 176 for instance, for the phase-processes that can be undertaken by material objects.

identity.[50] As for the commonly regarded 'laws of co-instantiation', for Bird these are not laws at all – the fact that certain properties are non-accidentally co-instantiated does not require the positing of substantive universals (and other attending laws), but can simply be seen as the result of the causal interactions governed by causal laws.[51] Thus Bird says that 'Water is H_2O', for instance, is not a law at all. The co-instantiation of the properties of the H_2O molecules (just like, presumably, the co-instantiation of water's superficial properties) is the result of other causal laws.

In fact, this view of the unity of the co-instantiated properties is the basis for Bird's proposal regarding the metaphysics of natural kinds. Kind members do not instantiate substantive universals but possess (directly) clusters of properties whose unity is the result of the operation of causal powers. Bird thus adheres to criterion VI_a regarding the patterns of properties. Kinds could be represented as points in the logical space of properties (when the laws produce a perfect agglutination of properties, possessed by *all* kind members) or as regions of density in this space (when the laws produce an imperfect agglutination so that the kind members possess most of the properties of the patterns characterising the respective kinds). Obviously, criteria I-III, concerning the nature of *properties* that are part of the respective patterns, are adopted by Bird as well.

Richard Boyd is the philosopher who firstly proposed that we should see the co-instantiation of properties by kind members as the *result* of something other than essences/substantive universals. In fact, Boyd's main concern is with the N&S requirement (criterion VI_a) and the fact that in biology organisms subject to classifications do not present clear-cut patterns of properties (i.e. do not seem to fulfil any necessary and sufficient conditions). His solution is to drop requirement VI_b and stick to VI_a (kind members in biology should instantiate clusters of properties) as well as to appeal to *mechanisms* as the external means through which the 'causal structure of the

[50] Bird construes powers in Ellis's (rather than Lowe's) style. For an individual, to instantiate such a non-substantive universal is to instantiate powers that produce, following causal interactions in the same or other individuals, the instantiation of other properties (the non-substantive universals that are in an identity-bearing relation with the initially mentioned universal), provided that certain conditions (e.g. the absence of antidotes) are in place. Powers are taken then to be intrinsic properties that, just like in Ellis, *can* produce the instantiation of other properties, where this *modal* level is worked out in counterfactual terms involving particulars and their properties (Bird, 2005: 437-428).

[51] *Cf.* Bird (2007: 208-211)

world' produces the agglutination of the properties characterising these (biological) kinds.

These mechanisms, which can be 'intrinsic' to organisms (e.g. developmental mechanisms) or 'extrinsic' to organisms (e.g. mechanisms involved in ecological adaptation, gene flow and interbreeding, evolutionary pressures), maintain the homeostasis of organisms. They are shared by the members of a population and are responsible for the resulting (clusters of) phenotypic and genetic similarities. Members of a biological kind need to possess most of these clustered properties and even if no 'laws of co-instantiation are present',[52] mechanisms and the attending causal interactions are sufficient to make such kinds 'inductively rich'.

In other words, Boyd adheres also to criteria I, II and V. Notably, mechanisms are supposed to agglutinate the determining properties (e.g. genetic ones) in connection with criterion VI_a, as well as to intercede the causal interactions between the determining properties and the 'superficial' ones (e.g. morphological features), in connection with criterion II.[53]

Just like Boyd, *John Dupré* focuses on biology and offers a very liberal account of kinds, in line with his general view that the world, in spite of scientists' unifying efforts, is disordered and messy. We could not discover any necessary and sufficient properties shared by categories of organisms, says Dupré (no matter how hard we looked at intrinsic similarities like the genetic or phenotypic ones, or relational properties like ecological

---

[52] ...even if no essences of biological kinds are in place, says Boyd (see Boyd, 1991: 129). In addition, for Boyd, mechanisms make kinds inductively rich, in spite of the fact that in biology we do not have 'causal' universal generalisations. Boyd writes that '[W]hat the [HPC] theory of natural kinds helps to explain, is how we are able to identify *causally sustained regularities* that go beyond actually available data and how we are able to offer accurate causal explanations of particular phenomena and of such causally sustained regularities (italics original, Boyd (1999: 152), *apud* Tsou, (2008: 72)). Boyd continues '[S]cientific (and historical and everyday) knowledge often depends on our being able to identify causally sustained generalizations that are neither eternal, nor a-historical, nor without exceptions, and our ability to do so depends on our coordination between language and classificatory categories and causal phenomena involving and defined by imperfect property homeostasis' (*ibid,* p. 164).

[53] In addition, the homeostatic nature of those mechanisms makes it that the *prima facie* 'profound' properties might also depend on the 'superficial'. This may happen due to the feedback interactions during organisms' lives and to the selection pressures, which make the instantiation of genetic properties depend on the phenotypic traits, across generations. See Boyd (1999:165) on the involvement of his homeostatic kinds in evolution and also Griffiths (1997) who develops Boyd's account and stresses the importance of extrinsic mechanisms for the homeostatic kinds to be found in the special sciences.

niche sharing) and our biological classifications do criss-cross each other.[54] Dupré thus denies criteria VI_b and VII and draws radical conclusions from this biological lesson. He advocates pluralistic classifications that are fuzzy, nested and depend on our 'epistemic' interests down to the point where similarities that are 'of interest for theoretical reasons' are no more fundamental than 'economically useful or strikingly noticeable' ones. Gastronomic classifications, for example, are considered just as good as, say, the ones framed in evolutionary biology.[55]

Notably, while maintaining that classification depends on our interests, Dupré does not allow classification based on mind-dependent, Cambridge properties.[56] Dupré therefore adheres to criterion I and feels entitled to call his position '(promiscuous) realism'. In his more recent work, Dupré even seems to accept criterion II by suggesting that the properties in question should after all provide a basis for scientific theorising.[57]

In this section I have presented the views of some of the most important theorists of kinds, who weigh differently the importance of the general criteria of kind membership (see fig. 3 above). They look at different scientific domains and explore various ways in which the kinds discussion is related to other debates in metaphysics and the philosophy of science. Their accounts will play a central role in my enquiry. Even if I lean towards considering some of these views as more robust than other, I will try to remain neutral in my own enquiry, in a methodological sense which will be discussed in the final section of this chapter.

In the next section I will outline some of the approaches to medical kinds that have been sketched in the literature, either by choosing some subset of the criteria for kind membership described in §1.1, or by adopting one particular account of natural kinds or another.

---

[54] Dupré (1993: 18), Thalos (1995: 252)
[55] Dupré (1981: 82-83) Dupré mentions as well the mapping of kinds in the multi-dimensional property space, leaving indeterminate what sort of properties should be reckoned with in this mapping.
[56] See Dupré (1993: 105) and also his response to Cooper's charge of arbitrariness in Cooper (2005: 50)
[57] *Cf.* Dupré (2002) *apud* Cooper (2005). At any rate, Cooper does develop a variant of his account in which the properties used in classifications are determining, explanatory powerful properties. See Cooper (2005: 50-52)

§1.3 KINDS IN MEDICINE

A number of philosophers and social science theorists have approached the theme of natural kinds in medicine (especially psychiatry) by appealing directly to various membership criteria or by adopting particular accounts of kinds. In the following, I will present the views of some of them that reflect the state of the art in the current disputes over medical kinds.[58]

Lawrie Reznek discusses medical kinds in a few chapters of his (1987) book dedicated to the nature of disease. The criteria Reznek adheres to are I, II, III and VI_a. Objects are said to belong to natural kinds if '(1) they have the same cluster properties (2) the possession of the cluster is explained in the same way (by the same laws of nature) and (3) natural boundaries exist between them and members of neighbouring kinds'.[59]

Reznek argues that *disease*, as such, does not represent a (genus) natural kind because different pathological conditions have different explanatory natures and the appeal to such explanatory natures does not help us in distinguishing diseases from non-diseases.[60] The border between the healthy and the pathological depends on our values. As Reznek puts it, this border 'is invented by us'.

However, particular pathological conditions could be natural kinds, since the organisms suffering from one disease or another exhibit a cluster of symptoms that might have the same explanatory nature. For Reznek, such explanatory natures should delineate the identities of diseases, or the identities of diseases as entities. More precisely, the identities of diseases should depend on the identities of the explanatory natures, where these explanatory natures are also clusters of properties that are distinct 'if and only if there is a relative rarity of intermediate forms' between them.[61] We might have troubles with those

---

[58] On inspection, some of their views are interesting and challenging. In general, I should point out from the start that their approaches reflect broad conceptual confusions and a very superficial level of the discussion. That is why in this section I shall extensively use quotations, for illustrative reasons.

[59] Reznek (1987:42) That this is what being a natural kind consists in is argued in chapter 2, in pp. 33-47. Reznek's most serious worry is that we might not be able to distinguish between the universal generalisations that are accidental and those that are not accidental (p. 41).

[60] Reznek (1987: 68). Reznek thinks that considering diseases as a natural kind is committing an 'essentialist fallacy' – attributing the same essential, explanatory natures to all disease cases. (*ibid.*, p. 78)

[61] Reznek (1987:41)

pathological 'entities' that vary continuously from the 'invented' healthy state. We might also have problems in deciding whether diseases which share only 'a part' of their explanatory natures are distinct 'entities' or are just forms of a more general, 'syndrome' entity. But overall, there is some 'grain of truth', says Reznek, in the assumption that natural kinds are to be found in the medical realm.[62]

Rachel Cooper dedicates a chapter to natural kinds in her (2005) book concerned with classification in psychiatry. Cooper makes the distinction between two important senses of essential properties – explanatory powerful and part of necessary and sufficient conditions of membership,[63] and notes in addition on the involvement of natural kinds in laws, explanations and inductions. The N&S requirement is dropped after espousing Dupré's convincing critique of its application on the biological realm and is substituted with the cluster requirement.

Cooper takes up criteria I, II, III and VI_b and investigates whether psychiatric disorders are natural kinds by using the mapping of kinds into the multi-dimensional space whose coordinates are represented by determining properties (*apud* Dupré). Cooper argues that in this space, psychiatric disorders as natural kinds could be identified with the regions of density corresponding to the psychiatric cases that at any stage of their illnesses present the symptomatic characteristics outlined by the conventional psychiatric categories. Not all diseases will qualify as natural kinds. There will be some psychiatric categories we can clearly map in this space (e.g. Huntington's chorea, whose genetic basis we know). But there will also be intermediate, doubtful cases (e.g. schizophrenia, whose symptoms, researchers suggest, do not share a sufficiently similar determining basis) and 'rag bag' cases that could not be natural kinds (like the category of Sexual Disorder Not Otherwise Specified).[64]

Just like Reznek, Cooper points out that the value involvement of the disease concept

---

[62] Reznek, (1987: 188, 9)

[63] Cooper (2005: 47, 58, 59). For clarity, Cooper calls the explanatory powerful properties *determining* properties, and as I have already noted, I adopt this use as well.

[64] *Cf.* Cooper (2005: 71-74) and Cooper (2004: 130). Cooper allows that the determining properties could be relational in the case of the diseases whose environment-dependence is decisive (Cooper 2005: 51). Other criticisms advanced against psychiatric diseases as natural kinds (e.g., Hacking's looping kinds argument, McGinn's argument from multiple realisation) are also discussed and rejected.

should not deter us from considering diseases as (candidate) natural kinds. It may be that a value judgement is involved in grouping together all the cases that medicine deals with under the heading *disease* (and indeed, Cooper argues that a value component is involved in the disease concept), but at least some of the cases that are treated, researched, etc. could qualify as natural kinds. A useful analogy drawn by Cooper is with the concept of weed – it might be that considering a plant a weed is a value judgment, but that being said, the plants put under this heading (e.g. daisies, buttercups) could be investigated as natural kinds.

Peter Zachar does not think that psychiatric disorders are natural kinds. Natural kinds, argues Zachar, are 'exclusively defined with reference to inherent properties'. They could be found in psychiatry only if diseases were 'bounded entities in nature' whose characteristics could be exhaustively circumscribed by statistical means (at the level of symptoms) and/or by the study of 'biopathological processes' (on the microstructural level).[65] This view of diseases is characterised as 'essentialist' and is taken by Zachar to be inconsistent with the emphasis on relational properties and the environment dependency that is present in psychiatric practice and the contemporary construal of biological species. Zachar thus argues that 'thinking in an anti-essentialist fashion and conceptualising psychiatric disorders as practical kinds [i.e. as relatively stable patterns of symptoms] is more consistent with a scientific view of the world'.[66]

Nick Haslam disagrees with Zacher. Haslam's objection is that Zachar is too narrow in his approach and that, by thinking in an anti-essentialist fashion, he ignores the possibility of property clusters that possess coherence and crisp boundaries:

'One problem with Zachar's analysis is that he identifies natural kinds with essentialist categories. Although there is ample precedent for this equation (e.g., Kripke), it is arguably possible to separate essentialism from the natural kind concept (Haslam, 1998). Boyd (1984), for instance, has proposed an understanding of natural kinds as homeostatic property clusters that possess coherence and crisp boundaries but lack necessary features. Most biological kinds are probably well understood in this fashion. It is well known that species cannot, after Darwin, be seen in an essentialist manner, as they lack necessary features, maintain their boundaries dynamically and relationally, change over time, and occasionally show gradations between

[65] *Cf.* Zachar (2000: 167)
[66] See Zachar (2000: 167) and also Benjamins (2003) who comments on Zachar's work.

species. However, they remain natural kinds in any meaningful sense of the term. (Haslam, 2003:3)

Having portrayed his alternative account of natural kinds – namely Boyd's, Haslam is quick to point out where Zachar's analysis fails – even if no essences of psychiatric disorders exist, it is not impossible that we could cut nature at its joints and attribute a respectable ontological status to mental diseases. In his own approach, Haslam suggests that psychiatric diseases could be natural kinds insofar as 'discrete categories, based on objective discontinuities rather than pragmatic decisions, exist in the psychopathological domain'.[67]

When we ask whether a mental disorder is well understood as a natural kind, I would argue, we are asking not whether it has an essence but whether it has a comparable ontological status to naturally existing biological kinds. That is, we are asking whether it has relatively crisp boundaries, is objectively real rather than merely reflecting classificatory conventions, and has a biological basis. Are mental disorders akin to biological species, with an equally solid, non-artifactual foundation in the natural world? In many or most cases, the answer will be a clear "no" However, the answer is obviously and uninterestingly "no" if we require a natural kind to have an essentialist basis that is implausible even for biological species. Almost no-one doubts that "carving nature at the joints" -- even if it is an ugly metaphor -- is a reasonable and feasible goal when it comes to biological taxonomy at the species level, even if these joints are not defined by essences. The question ought to be whether we can have the same confidence in the existence of "natural" joints and objective discontinuities in the psychiatric domain. In short, I think Zachar deflects the real challenge of the natural kind view of mental disorders by joining it to the important but distinct issue of essentialism.' (Haslam, 2003:3)

Daniel Sulmasy discusses the connection between natural kinds and diseases in his (2005) article, which is meant to espouse an Aristotelian position. For Sulmasy, natural kinds must be attended by law-like principles and their members must possess 'dispositions'.[68] Diseases are always diseases of a living natural kind but they are not themselves natural kinds because they fail to fulfil a crucial criterion – the dispositions of a natural kind should allow any one of its members 'to flourish as the kind of thing that it is'.[69] This criterion is unfolded as follows:

'Natural kinds have dispositions. This much teleology must be granted. Uranium undergoes a characteristic

[67] Haslam (2003: 4) See also Tsou (2008) who adapts Boyd's account to psychiatric kinds.
[68] Sulmasy (2005: 491, 492)
[69] Sulmasy (2005: 492)

pattern of radioactive decay. Various types of stars have dispositional predicates. They develop and change over periods of time that may seem long by human standards, but there is a pattern by which a star's history must unfold temporally if it is to behave as the kind of thing that it is. Dispositional predicates seem especially characteristic of living natural kinds. In fact, Philippa Foot has written that the word 'good' is an attributive adjective, not a predicative adjective, and cannot be understood apart from an understanding of what kind of thing something is. So, for instance, the word 'good', as used in the phrase, "good roots," cannot be understood unless one knows that it is being attributed to a rosemary bush and not to a rhinoceros. Having deep roots would not be good for a rhinoceros. Natural kinds have natural tendencies. Much of the scientific enterprise consists of coming better to understand these natural tendencies. This teleology....does not imply the anthropomorphizing of things...[T]hat natural kinds have law-like principles that determine how they develop and flourish as the kinds of things that they are is simply a fact about the world as we encounter it.' (Sulmasy, 2005: 492-493)

Since diseases fail to contribute to flourishing, they are not natural kinds but classifications 'of a certain state of affairs that can occur in members of particular living natural kinds'. I note en passant that Sulmasy mentions several other aspects of this construal of diseases, the asymptomatic aspect being one of them.

'There can be asymptomatic disease. But if a pattern of disturbance in the law-like biological principles that determine the characteristic development and typical history of a living natural kind is to be called a disease, at least some individuals with the disease must be inhibited from flourishing as the kinds of things that they are. For example, prostate cancer at age 80 may be "incidental" and never interfere with a man's flourishing. But unless prostate cancer interfered with at least some men's flourishing, it would not be called a disease.' (Sulmasy, 2005: 497)

In his (2006) book dedicated to psychiatric classification, *Dominic Murphy* argues that it does not matter whether (psychiatric) diseases are natural kinds or not.[70] Murphy notes first the variety of ways in which natural kinds have been construed and organises the accounts on offer in four categories: simple essentialisms, nomological kinds, refined essentialism and ecumenism. Simple essentialism is 'the view that what makes a collection of entities into a natural kind is some microstructural essence that accounts for their other properties'.[71] The nomological view states that kinds participate in the laws of nature; in this approach, since the laws of nature are 'the deepest, immutable regularities that there are', the involvement of natural kinds in laws entails that they should also be

[70] Murphy (2006, 9.4, 9.5.5)
[71] Murphy (2006, 9.5.1)

regarded as 'spatiotemporally unrestricted and immutable'.[72] The 'refined essentialism' is represented by Boyd's account, which preserves the core of simple essentialism in that 'microstructure is one way to keep properties together', but has the additional insight that due to mechanisms, 'clusters of properties may stick together homeostatically, despite perturbations in the environment'.[73] Finally, ecumenism is the variant of Boyd's account developed by Griffiths, which identifies inductively useful, homeostatic clusters in other special sciences beside biology and stresses the importance of external mechanisms.

Murphy allows himself for a moment to be drawn into the debate and expresses his discontentment with ecumenism. The main reason seems to be that, by stressing the importance of external mechanisms, Griffiths deviates from the refined essentialist line represented by Boyd's account.

'Griffiths thinks that this liberalisation of the notion of essence is justified by its resonance with Boyd's ideas and by its continuity with earlier theories of natural kinds (1997, 189). The idea that the essence of something can be external to it was never part of the tradition of natural kinds, but it is not ruled out by the idea of homeostatic mechanisms to which Boyd appeals. But if extrinsic forces are permitted to hold a kind together, then maybe chartered accountants are a natural kind, since they share properties in virtue of a historical fact – they passed the relevant exams. But now we start to wonder whether anything can be a natural kind' (Murphy, 2006, 9.5.5)

The problematic tension between refined essentialism and ecumenism appears to have deeper roots, down into the preferences that each theorist might have. Evidently, a natural question arises – why should we care?

Why should we care though? Why not break with tradition, do away with essences altogether, and embrace ecumenism? The great insight of ecumenism is the recognition of the many reasons for properties to cluster together, ranging from atomic structure to government declaration. So both refined essentialism and ecumenism are good concepts of kindhood. My taste runs to saying that the former picks out natural kinds and the latter does not, although it does pick out useful kinds. But I do not have a knockout argument against all other positions. It is hard to see how to resolve the dispute over whether all kinds are natural, or only ones with simple essences are but that there are also other kinds of kinds.....Whether or not we call some or all of these natural kinds does not seem terribly consequential. What is the significance of this for psychiatry?' (Murphy, 2006, 9.5.5)

---

[72] Murphy (2006, 9.5.2)
[73] Murphy (2006, 9.5.4)

Since different theorists might have different opinions as to what natural kinds are, whether diseases are natural kinds or not is in fact a moot point, as 'mental illnesses are kinds, even if they are not natural kinds. And whether they are seen as natural kinds 'is largely a matter of choice'.[74] Murphy's conclusion is that we should just look at useful ways of classifying diseases (for treatment) and leave aside their natural kind status, since 'there are both natural kinds and other scientifically fruitful divisions in nature'.[75]

The views of the theorists presented in this section are representative for how natural kinds have been discussed vis-à-vis the medical domain. Some of these views are interesting and challenging. And yet, I tend to think that they reflect a problematic relation to the very issue of what natural kinds are – or, in other words, to the realism vs. nominalism dispute concerning kinds – simply because certain assumptions about kind realism are taken for granted while other assumptions are not even mentioned. I will discuss this realism dilemma in the next section, in which I will also state the path of my subsequent enquiry.

To anticipate, what I will try to do in the next chapters is simply to complement the argumentation of those theorists who claim that medical kinds fit coherently into a picture of natural kinds (especially 'cluster' kinds) by taking into account as many criteria of membership as possible.

[74] Murphy (2006, 9.5)
[75] Murphy (2006, 9.5)

§1.4 Realism (and Nominalism) about Kinds

What is natural kind Realism and how can we define it? When investigating the status of kinds of diseases (or any other kinds in the special sciences), the issue of kind Realism (and Nominalism) becomes almost a methodological concern. The reason is that the discussions about the ontology of natural kinds can involve extremely difficult and variegated metaphysical arguments, which are mostly set out against the background of the exact sciences. When entering the relatively uncharted territory of the special sciences then, there is not enough scope to discuss fastidiously enough about all the ramifications of the kind Realism debates. Accordingly, there is a problem as to how one should connect one's enquiry into the special science kinds to the classical accounts of natural kinds.

There exist two extreme strategies that a theorist of the special sciences could support, corresponding to two methodological 'dangers', as it were. On the one hand, the 'danger' when discussing a science as medicine is simply to take up one account of natural kinds or another and 'translate' the results into the respective special science, drawing conclusions much too swiftly. This does happen in most of the cases of the authors discussed in the previous section. On the other hand, the 'danger' is to insist upon the variety of kind membership criteria and adopt the position that 'natural kind' is an indeterminate and useless notion, which one can employ in whatever sense one wishes to.[76]

I do not think that either of these extreme positions is fruitful when dealing with the issue of kind Realism. I will state here my reasons and then portray the 'middle ground' methodology I will take up in my enquiry. As far as the first extreme is concerned, a theorist of special sciences should not ignore that there are several options available for the kind Realism/Nominalism dilemma and that none of them has proven to be beyond (reasonable) criticism. A simple way to present some of these options in a unitary way goes as follows.

---

[76] Murphy (2006) adopts this position. For another example see Machery (2006) - who discusses the natural kinds of concepts and notes, as a pre-emptive defence of his preferred account, that 'the notion of natural kind, like the notion of heap, is vague' (Machery, 2006: 448). See also Shain (1993: 291) for a similar stance.

A very strong option is to take kinds as real only insofar as substantive universals exist.[77] In this respect, one could argue that the entire debate over kind Realism hinges on what ontological status we should assign to the patterns of co-instantiated properties we find in nature. If, ontologically (from the point of view of a classical substance-attribute relation), the co-instantiation of certain properties by individuals does not represent anything more than 'the sum' of separate instantiations of properties (taken singly) then one can only adopt (some form of) reductionist stance on kinds. If substantive universals exist, then the phenomenon of co-instantiation calls for a separate metaphysical category. Otherwise, all criteria underlying the various ways in which our classifications are non-arbitrary will be related to forms of kind Nominalism. For instance, denying criterion I (natural properties requirement) would amount to adopting a strong nominalist position, affirming criterion II (causal induction requirement) would be an allegiance to moderate nominalism, and, say, embracing criterion VIb (N&S requirement) could be viewed as adopting a weak nominalist stance.

This approach might appear too radical and one might want to shift towards a less ambitious option. For instance, it is somewhat plausible to regard the position that attends criterion VIb (N&S requirement) as realism or weak realism.[78] If the patterns of properties instantiated by particulars are clear cut (as opposed to being fuzzy), then one could argue that kinds do represent a separate category that delineates a special place in our ontology - and the aesthetic, hierarchical arrangement of chemical elements that are separated by gaps certainly invites this line of reasoning. Intuitively, kinds are some sort of 'entities' and having non-fuzzy patterns, as opposed to fuzzy ones, seems to go sufficiently far in this (intuitive) direction. We simply do not need substantive universals for kind Realism, one could then claim.

However, in fact, one could find reasons to claim that each of the criteria outlined in §1.1 is sufficient to delineate kind realism. One could accept for instance that when particulars instantiate several explanatory powerful properties, we have (from the point of view of a substance-attribute relation) a 'sum' of 'separate' instantiations, insofar as no substantive

[77] It is for instance Lowe's (2001) position. Lowe does not distinguish between variants of nominalism, as I shall do in the following.
[78] Ereshefsky (1992) (1998) for instance, holds that it is a realist position

universal exists to account from the above for the phenomenon of co-instantiation. Nevertheless, one could point out that the 'sum' represents ontologically something more than its 'parts'. As I have underlined in §1.2, it has been pointed out that objects sharing a single (explanatory powerful) property (e.g. charged bodies), do not seem to form a natural kind. Not just properties, but patterns of properties are instantiated and these patterns (irrespective of whether they are fuzzy or not) should delineate a separate metaphysical category. One could thus give up criterion VIb and instead of necessary and sufficient conditions, simply ask that clusters be in place, embracing criterion VIa. 'Insofar as clusters of properties exist, then kinds are real', the conclusion of this approach would sound.

One could go all the way down and even take just criterion I as sufficient for kind realism.[79] The point would be that there is still an important difference between classifications based on Cambridge properties and the ones based on natural properties. In addition, it could be argued, unless individual substances are construed as a 'bunch of properties', as most often the trope theorists hold, then individuals instantiating a property and grouped on grounds of this similarity do appear to mark out a separate form of being.

The variety of these options is indeed striking and the discussion of which of these options (or any other I have not portrayed here) points towards Realism is a very complicated matter. However, turning to the other extreme, I do not think that one should adopt the position that natural kinds represent a 'dummy', indeterminate notion that everyone can use in whatever sense one wishes. Indeed, sometimes the special science theorists point out that scientific, empirical enquiry can dispense with the notion of natural kind.[80] This may be a correct point - perhaps for medical research it does not matter much whether there are natural kinds of diseased organisms - but it does seem to miss the target, simply because from the point of view of scientific practice only, many current philosophical discussions may appear redundant. From scientists' point of view, a

[79] I have already mentioned (§1.2) that Dupré takes his own position as a 'promiscuous realism', because the properties involved in his most unusual classifications, are not gerrymandered. See Wilson (1996), Daly (1996), Wilkerson (1998) for some samples of the feisty debate on whether this deserves the tag of realism.
[80] This is Murphy's (2006) stance, for instance.

discussion such as whether laws are metaphysically necessary or not, for instance, may appear quite futile as well, since they could of course make do with the notion of physical necessity only. To give another example, scientists arguably do not care much about (and their research is not influenced by) the dispute on whether properties are universals, tropes or sets, and the examples could go on.[81]

Nonetheless, all talk regarding the thorny relationship between empirical research and theoretical reasoning notwithstanding, 'natural kind' does represent an important metaphysical notion found at the intersection of various discussions in metaphysics and in the philosophy of science, which inter alia can be used to address the distance between the exact sciences and the 'immature' ones. Therefore, discussing natural kinds is useful and an investigation of their status in medicine can be fruitful, if conducted on the right track.

As I said, in order to avoid the two extreme positions, I will adopt a simple, 'middle line' methodology. On the 'negative' side, my approach presupposes, for one thing, that I will avoid getting involved in the kind Realism debate and/or avoid trying to adjudicate between various particular accounts of natural kinds. For another, it entails that I will resist 'projecting' directly the conclusions of any particular natural kind account from the exact sciences into the realm of the pathological.

On the 'positive' side, my approach will be to take each criterion associated with natural kind membership and see whether an 'ontological gap' between the exact science kinds and the medical kinds opens up. More precisely, in chapter 3 I will compare a 'classical' natural kind (namely gold) with a kind of disease (namely Graves' Disease). The exact science kinds are the ones considered to typify natural kinds and I take it as uncontroversial that, irrespective of what resolution the Realism dilemma over the right membership criteria might have, if natural kinds exist, they should be found primarily in the exact sciences. Such a comparison should thus make clear certain resemblances and differences between medicine and the exact sciences that should matter decisively in establishing the character of disease kinds.

---

[81] Of course, philosophical arguments *could* be brought in order to show the redundancy of the natural kind notion. The most important one is advanced by Ian Hacking in his (2006) and (2007).

My 'middle line' methodology will not be entirely neutral, however, in the sense in which I will not simply use these criteria as means of comparison without questioning their grounding in the intuitive backdrop of the natural kind discourse. As I have explained, this intuitive backdrop poses a great challenge to the intelligibility of my enquiry and its relation to the kind membership requirements needs to be explored. I will explain how in the next subsection.

§1.5 CARVING NATURE AT ITS JOINTS

Intuitively, natural kinds are supposed to point towards classifications that are 'profound', 'ultimate', 'basic', 'fundamental' and reveal 'what things are'. Of all the expressions of these intuitions, I shall take the idea that natural kinds 'carve nature at its joints' to be representative.

Kind theorists often appeal to this phrase when they want to make clear that their preferred account of natural kinds is the right one.[82] Nevertheless, 'what things are' and accordingly, what 'carving nature at its joints' amounts to, is not univocal. As such, invoking this assumption seems *prima facie* just a rhetorical device that comes to accompany whatever stance theorists adopt on the Realism/Nominalism dilemma discussed in §1.4. That is, its purpose simply seems to add some un-argumentative weight to one's claims that certain criteria of membership should be employed for delineating natural kinds.

This would be, however, too quick a dismissal of this assumption or set of assumptions. We do have the intuition that natural kinds pick out something 'deep', 'profound', 'fundamental', etc. While arguments based exclusively on intuitions are weak, intuitions as such cannot be altogether rejected and their importance should not be neglected. In the context of my proposed comparison, the respective intuitions tell us that Graves' Disease, like all disease kinds, is a 'superficial' kind and that comparing it with the 'fundamental' kind gold would be entirely inappropriate. What we have here is a

[82] See for instance Wilkerson (1995)

challenge to the very intelligibility of the present enquiry. How could this challenge be dealt with?

Again, I do not wish to reject these intuitions. Being aware of their rhetorical potential, I would rather ask - what direction(s) could they be leading towards, when revealing 'what things are' is at stake? I will outline two alternatives as to what 'carving nature at its joints' points to. Then I will introduce a caveat I already announced in the Introduction to this thesis and I will explain how, in spite of this obstacle to intelligibility, my enquiry in the next chapters is to be undertaken.

In contemporary metaphysics, there is a divide - among the possible metaphysical schemes that could be employed in order to substantiate the abovementioned intuition(s) - between two basic approaches on how revealing kinds is disclosing 'what things are'. In very simple terms, on one approach, being a member of a natural kind is crucial for the identity of any particular.[83] On the other approach, kind membership is not taken to be crucial for the identity of particulars.[84] Let me exemplify with two theorists who explicitly state in their work how they take kinds to 'carve nature at its joints' from this point of view, namely Brian Ellis in his work on laws of nature and David Wiggins in his work on sortals.

Like all sortal theorists, Wiggins takes sortal concepts to be associated with semantic conditions of 'identity and individuation' that 'sort out' which things fall under them and which do not. In his approach, however, these semantic conditions do not only allow the users of sortals to count things (an obvious consequence of the capacity to identify the objects falling under them)[85] but also, in the case of a certain range of sortals called 'ultimate', to establish the truth conditions of *a posteriori* (absolute) identity statements. More specifically, for Wiggins, statements such as 'a is identical with b' are true if and only if a and b fall under the same ultimate sortal f and they have the same causal

[83] Brody (1974, 1981), Denkel (1996), Wiggins (1980); see also Bigelow (1999)
[84] See Ellis (2001, 2002), Wilkerson (1995), Lowe (1989, 2006)
[85] In the traditional sortals discussion, the semantic conditions of identity should allow us to recognize and count objects. They are associated with relative identity, that is, with objects' identity *as* objects of a sort or another. See Geach (1962) and also Noonan (2006)

history.[86]

What this means on a metaphysical level is that, should any object a cease to fall under an ultimate sortal f, it would cease to exist. In other words, the possession of the characteristics that objects 'of a sort' share – characteristics also involved in the semantic conditions of application for sortals - constitutes metaphysically a *sine qua non* condition for preserving the identity and existence of the objects in question.

Converted into the kinds terminology, Wiggins's contention is that the possession of certain properties that objects 'of a kind' share in virtue of their membership constitutes metaphysically a necessary condition for the identity and existence of kind members. Were natural kind members to lose their kind characteristics, they would cease to exist. Wiggins holds, for instance, that:

'...the *elucidation* of the identity 'a=b' depends on the **kind of thing** that a and b are...Consolidating all this we have a threefold theory...a=b if and only if there exists a sortal concept f such that (1) a and b belong to a kind which is the extension of f (2) to say that x falls under f – or that x is an f – is to say what x is (in the sense Aristotle isolated) (3) a is the same f as b, or a coincides with b under f, i.e. coincides with b in the manner of coincidence required for members of f (Wiggins 1980: 48, italic original, bold added)

'[A]n individual *a's having ceased to exist at t* is a matter of nothing identical with a belonging to the extension of the ultimate individuative k that is a's kind' - Wiggins 1980: 67, italics added)

Now, Brian Ellis exemplifies the other branch of the divide. Ellis also thinks that kinds 'carve nature at its joints' and reveal 'what things are'.[87] In his rationale however, this role played by kinds has nothing to do with the individual identity of kind members. As I have explained in §1.2, in Ellis' view, the similarities used to demarcate kinds are dispositional properties and microstructures selected precisely (and only) because they are explanatory powerful. In his approach, membership to a kind brings 'what things are' into the light in the sense of explaining the behaviours that kind members exhibit. Only properties and their causal interconnectedness are in focus.

---

[86] In other words, Wiggins adopts the view that semantically, sortals have associated 'absolute' and not only relative conditions of identity. I should emphasize that in the present section as well as in chapters 2 and 3, the term 'identity' is chiefly used to refer to absolute identity. *Relative* identity will only be brought into discussion in §3.6, in a clearly specified context.

[87] e.g. Ellis (2000: 335)

Accordingly, the (individual) identity of kind members is eluded from the rationale underlying Ellis's affirmation that 'natural kinds carve nature at its joints' (on his preferred interpretation of that phrase). Parenthetically, Ellis uses at times the phrase 'identity of things' indistinctly, in that it is not clear whether in his argumentation is at work the absolute identity of particular kind members (*qua* particulars) or their relative identity *qua* kind members; this equivocation reverberates in some of his other arguments on kinds, which will be looked at in the next chapter. An extended survey of his work is revelatory though – for Ellis, kind membership has nothing to do with the absolute identity of kind members, as the following quotes indicate.

'The individual essence of a thing is the set of its characteristics in virtue of which it is the *individual* it is...The *kind* essence of a thing on the other hand, is the set of its characteristics in virtue of which it is a thing of the kind it is. Of these two conceptions of essence, the most important one...is the kind essence... (Ellis 2002: 12)

Things behave as they do...not because of any external constraints that force them to, but because this is how they are intrinsically disposed to behave in the circumstances...I assume that the identities of thing of the most fundamental kinds are wholly dependent on how they are disposed to act. If protons, for example, are such fundamental things, then for anything to be a proton it must always be disposed to behave as protons...protons could not behave according to different laws without ceasing to be things of the kind they are' (Ellis 2002: 2-5)

I am reluctant to accept...that the individual essences of a thing belonging to a natural kind includes its kind essence. Nor do I think that I have to accept this thesis to provide a sound basis for scientific essentialism.' (Ellis 2002: 238, italics original)

We see that, on one approach, illustrated by Wiggins, carving nature at its joints via the kinds scheme is in principle providing (a part of) the conditions of identity for individuals. On the other approach, illustrated by Ellis, nature is just carved in a different sense. Each of the two resulting 'branches' admits various detailed interpretations, specifying additional metaphysical characteristics that kinds are taken to have and consequently articulating the 'being' of kind members (i.e. 'what they are' in virtue of their membership). For Ellis for instance, as I have outlined in chapter 1, natural kinds are characterised by 'essential' properties and kind members instantiate substantive universals arranged in a hierarchical structure. For Wiggins, members of any natural kind

share a certain 'principle of activity', which is 'nomologically grounded' in causally efficient properties discovered *a posteriori*.[88]

I cannot claim that this divide is the very crucial one for elucidating the crux of intuitions attending the natural kinds theme. In other words, I can accept that one may attempt to tackle in a different way the substance of the intuition(s) in question.[89] Nonetheless, for my purposes it is sufficient that undeniably, this is an important divide. In its light, my enquiry in the next two chapters can be plausibly undertaken, in spite of the problems of intelligibility I have portrayed in this section. The guiding rationale will be the following: natural kinds can be taken to carve nature at its joints down to the identity of their members or not. If identity is indeed involved in circumscribing a 'basic ontology', providing an 'ultimate classification', revealing 'what things are', etc., then there is no way diseases could be natural kinds. In other words, there is no way the identity of an organism could depend on whatever diseased traits it may acquire. This is the very caveat of my enterprise, which I had mentioned in the Introduction.

However, if identity is *not* involved, then for every metaphysical feature that is present in the case of the gold kind, we could find a correspondent in the case of the Graves' Disease kind. Whatever 'essential' properties are said to be possessed by the members of the gold kind, similar 'essential' properties could be said to be possessed by the members of the Graves' Disease kind. If a substantive universal was said to be instantiated by the gold kind members, then I will claim that there is no reason why the Graves' Disease kind members should not also instantiate a Graves' Disease substantive universal, and so on.

I will follow the latter side of my rationale, in which identity is not explicitly involved, in chapter 3, by using the criteria of membership as means of comparison between the Graves' Disease kind and the gold kind. Prior to that, however, some of these criteria of membership (namely requirements VIb, VII and VIII) will be scrutinized in chapter 2, as part of a more general, metaphysical discussion of the identity-sense of carving nature at

---

[88] *Cf.* Wiggins (1980 : 77-83 *et passim*). See also Weatherson (2002) for a discussion.

[89] I could accept for instance that instantiating substantive universals or not is another important divide, as I think Lowe would claim (Lowe 2006). My divide, however, criss-crosses all the other differentiations, insofar as no major metaphysical aspect is neglected in my discussion. I do not ignore for instance that natural kind members could be said to instantiate substantive universals. Indeed, I shall discuss this aspect extensively in the next chapter. I am just trying in this section to deal with a set of intuitions.

its joints. These three requirements, I will argue, are related to the identity sense of carving nature, in that only if the identity of kind members depends upon their membership, are these requirements justified. Thus, I will show that the problems of essences, substantive universals and processes - i.e. the specific metaphysical problems we are confronted when questioning the justification of these requirements – are quite innocuous for the existence of medical kinds, if the identity sense of 'carving nature at its joints' is not explicitly invoked. [90]

As to the very justification of this identity sense of 'carving nature', in turn, I will show that it faces huge epistemological problems, the main reason being that the notion of identity is primitive.[91] I do not think that these epistemic difficulties can be overcome and I will offer at the end of chapter 3 some reasons for such scepticism. However, I will not draw a definite conclusion as to whether this identity sense of 'carving nature at its joints' is tenable or not.

[90] One consequence of this argument, in chapter 3, will be that to deny that the ontology of gold kind members is on the same metaphysical level with the ontology of Graves diseased organisms is to tacitly appeal to the individual identity of kind members.

[91] In the sense in which *about* identity, as a primitive relation entities bear with *themselves*, nothing informative can be said. See (Lowe 2001:44-47), Oderberg (2007: 117) and Mackie (1994). I will not discuss which of these two senses of carving nature is correct, from a *metaphysical* point of view; see § 3.6

CHAPTER 2 ESSENCES, SUBSTANTIVE UNIVERSALS AND PROCESSES

INTRODUCTION

This chapter is dedicated to an analysis of three criteria of natural kind membership - criteria VI_b, VII and VIII (the N&S, non-overlapping and non-phase requirements, respectively) – vis-à-vis the identity sense of carving nature at its joints. These criteria ask that (candidate) natural kind members should have necessary and sufficient conditions of membership (criterion VIb), that no two natural kinds could share members except when they are in a species-genus hierarchy (criterion VII) and that these members should not instantiate the kind specific-properties as part of a process in which more fundamental kinds are involved (criterion VIII). In other words, it is demanded that the 'divisions in nature' corresponding to natural kinds should be clearly delineated (non-fuzzy), non-overlapping (except in a genus-species hierarchy) and non-phase (non-implicated in transitory transformations).[92] What I want to argue is that that without invoking the diachronic identity of kind members, these requirements are groundless and that theorists appeal tacitly to the identity-related sense of carving nature when setting them forth.

My starting point in analysing criteria VIb and VII is the existence of determining properties and their patterns – the metaphysical backdrop that, as we have seen in §3.1, circumscribes the inductively rich kinds. Now, if we simply followed the patterns in which the determining properties are co-instantiated, we would find some 'divisions' that are fuzzy and criss-crossing, and other 'divisions' that are more clear-cut, as it were. Why should we consider the latter divisions ontologically higher than the former, in the pursuit for natural kinds? It is clear that, in order for the N&S and non-overlapping requirements to be justified, natural kinds should represent something more than groups

[92] In the notation I have introduced in chapter 1, criterion VIb says that, for any individual *a*, member of a kind K, the determining properties part of the pattern [p1, p2,...,pn] should be necessary and sufficient for membership. Criterion VII requires that, were *a* to also be a member of a different kind K*, then either K and K* are identical or they are situated in a species-genus hierarchy. Criterion VIII asks that the pattern [p1, p2,...,pn] should not represent a *phase* of a more fundamental pattern possessed by *a*.

of particulars sharing (patterns of) determining properties.

These requirements could be justified, by introducing new elements in the metaphysics of kinds, in two (more or less related) ways. One way appeals to 'essences' (and 'essential' properties) attached to kinds (or possessed by kind members). The other appeals to substantive universals said to be instantiated by kind members. I will argue, nonetheless, that what we end up with, when appealing to such metaphysical elements, is a chicken and egg dilemma. The dilemma goes as follows. When looking at the patterns of determining properties, should we pick up - among the 'candidate' divisions in nature that could correspond to natural kinds - only the non-overlapping and non-fuzzy ones because natural kind members are supposed to instantiate 'essential' properties and/or substantive universals? Or should natural kind members instantiate 'essential' properties and/or substantive universals because the divisions in nature corresponding to natural kinds are supposed to be non-overlapping and non-fuzzy?

I want to indicate in §2.1 and §2.2 that there is a vicious circle between, on the one hand, characterising in a certain way the metaphysical category of natural kinds (and the divisions of nature corresponding to them) and, on the other hand, attributing to natural kind members – which are already taken to (co-) instantiate (patterns of) determining properties, on a minimal interpretation - certain *other* types of instantiation. That the properties forming the patterns in question are not only determining but also 'essential' (on a certain construal of the essential) does not avoid the vicious circle, I will argue. Similarly, that the properties forming the patterns in question are not only 'determining' but their possession follows from the instantiation of a substantive universal, does not break the vicious circle. Of course, I will argue that the vicious circle can be broken down only if the identity-related sense of carving nature at its joints is appealed to.

My argumentation in §2.1 and §2.2 is intended to cover both criteria VIb and VII (N&S and non-overlapping requirements). Indeed, my point is that both of them are justified only if the identity of kind members is appealed to. However, for the ease of exposition, I will mainly discuss the N&S requirement and will come back to criterion VII whenever it is necessary to point out that the same treatment applies to it as well. To emphasize, everything that will have been said about criterion VIb is also perfectly relevant for the

non-overlapping requirement.

In §2.3 I turn to criterion VIII and the problem of the phase/non-phase distinction. I begin by accepting that intuitively, we have a certain range of kinds that appear to be *phase* kinds, i.e. appear to result from picking up (patterns of) properties that are only instantiated by particulars because they are engaged in certain (more or less transitory) processes. I ask, however, what is the *justification* for the commonsensical phase/non-phase distinction that underlies our usual dismissal of certain kinds as phases of more 'fundamental' kinds. I focus on what I consider the best philosophical defence in the literature of this commonsensical phase/non-phase distinction, namely Jonathan Lowe's, and argue that it fails in its current form and that its only hope is the adoption of the identity-sense of carving nature.

What is interesting about Lowe's position is that the diachronic conditions of identity for objects do play a role in defending the common-sense phase/non-phase distinction. These conditions are *not* linked with (non-phase) kind membership, however, but with more general categorizations of objects (in material, artefactual and biological strata). Lowe claims, among others, that unless these conditions of diachronic identity are accepted, the perspective of the number of natural, non-phase kinds multiplying greatly above the number of kinds commonly accepted (the so-called 'Heraclitan' threat) and/or reducing greatly below the number of kinds commonly accepted (the so-called 'Spinozan' threat) would be un-escapable. My point will simply be to show that these Heraclitan and Spinozan threats cannot be repelled by Lowe's strategy.

Of course, I will not want to suggest at all that it is meaningless to talk about phase kinds. I simply want to say that insofar as criterion VIII is used to rule out special science kinds as phase kinds, then the only chance to do so is by adopting the identity sense of carving nature at its joints. That the alternative of having special science kinds as non-phase kinds entails some form of 'Heraclitanism' does not represent in the least a catastrophic perspective. Such a possibility should simply make us re-think the phase, non-phase distinction, in other terms than the commonsensical ones.

§2. 1 ESSENCES AND ESSENTIAL PROPERTIES – THE LOCKEAN STRATEGY

One way of justifying why the patterns of determining properties should fulfil the N&S (and the non-overlapping) requirement(s) is to claim that the characteristic properties of kinds are not only determining but also 'essential'. In this line of justification, the necessary and sufficient conditions for membership should circumscribe 'essences' of natural kinds. Such 'essences' are evidently missing in the case of fuzzy (and/or overlapping) kinds, in which we only find patterns of determining properties *simpliciter*.[93]

I will show in § 2.1.1 that, for all the kinds that fulfil the requirement(s) VIb and VII, the 'essential' denomination attached to their characterizing properties is only a 'tag'. Just employing such a tag does not add anything to the nature of the properties in question (and hence to the nature of the divisions of nature they characterize). Such properties are ultimately determining properties. Hence, criteria VIb and VII are unjustified and their only justification, I argue in § 2.1.1, is the identity sense of carving nature.

I should stress that I do not intend to question (or even enter into the discussion of) whether kind members that do fulfil in the actual world certain N&S conditions would also have to fulfil them in all possible worlds. That debate concerns whether, for example, the division in nature that corresponds to the (natural) *kind* gold is metaphysically necessary circumscribed by a certain (non-fuzzy) set of properties. What I ask in this section is a different question: why *should* the divisions of nature (the groups of particulars with certain similarities) that correspond to *natural* kinds be so and so? In other words, my employment of the *'should'* modal level here is intended to question a certain characterization of the metaphysical category of *natural* kinds.

---

[93] e.g. Wilkerson (1995:61-88), Harré (1974) e.g. Wilkerson (1995:61-88), Harré (1974). Note that the N&S and/or the non-overlapping requirements do not *entail* the existence of 'essences' and 'essential' properties. 'Essences' could be brought in as a possible *justification* of these requirements, in a philosophical argumentation, and, as I have already mentioned, substantive universals represent a possible alternative justification. One could argue that, conversely, 'essential' properties do not entail these requirements, simply because 'essential' properties could be construed as being only *necessary* for membership (Bird, 2009). At any rate, as I have mentioned in chapter 1, the very term 'essential' has an extremely complicated semantics in that there are at least four senses of it that are invoked in the natural kinds discussion. What I will try to show in the following is that, in a sense, there is a relation of *de dicto* entailment between (a certain construal of) essential properties and the N&S and non-overlapping requirements and that the *de re-de dicto* slip hinges on an equivocation between certain senses of the essential.

Were natural kinds turn up *not* being characterized by criterion VI_b, this would be consistent, I believe, with arguing that the division of nature corresponding to gold is non-fuzzy in a metaphysically necessary (or only physically necessary) way, if the argument employed *causal* explanations for why certain properties are co-instantiated in perfect clusters.

That is to say, the argumentation that (the division of nature corresponding to) gold is metaphysically necessarily circumscribed by a certain set of properties, should not (uncritically) use as assumption(s) that

i) gold is a natural kind

ii) ii) *all natural kinds fulfil the N&S requirement*

iii) iii) any metaphysically possible world should possess the same structure of natural kinds as the actual world

iv) iv) in the actual world, gold is characterized by such and such properties

in order to derive the conclusion that

v) in any metaphysically possible word, gold fulfils the N&S requirement (and is characterized by such and such properties).[94]

Premise ii) needs to be argued in favour of *separately*, and not taken for granted, irrespective of *what is the case* with certain divisions of nature in the actual (or any possible) word.

§2.1.1 THE LOCKEAN STRATEGY

What are the 'essential' properties that supposed to form the necessary and sufficient conditions of membership for non-fuzzy kinds? When discussing them, the kind theorists embracing the N&S requirement appeal, ultimately, to the terms in which one would

[94] Parenthetically, premise iii) is employed by authors like Ellis (see Ellis *et. al.* 1992, Ellis 2008:146) who are interested in the metaphysical necessity of natural laws, and is quite disputable. Without premise iii) the conclusion that can be derived, at best, is that in any *physically* possible word, gold fulfils the N&S requirement (and is characterized by such and such properties); see Lange (2004) for a critical stance.

describe determining properties. That is, they are described as the properties that explain (i.e. produce) observable behaviours.

It should be said that the rationale employed by the kind theorists in question belongs to a certain Lockean tradition. Just as Locke, whom they cite or mention explicitly, these theorists insist upon the conventional nature of the classifications based on observable (secondary) properties. The claim is that the observable properties are mind-dependent and/or relational and thus (in my terminology) the resulting kinds fail to fulfil criterion I. The observable properties (assimilated more or less to Locke's secondary qualities) are contrasted with the micro-structural properties (which are taken to correspond to the Lockean primary qualities). The latter are adverted, on the one (rhetorical) hand, to be able to show 'what things are' and to 'carve nature at its joints', in virtue of their causal roles. On the other hand, they are said to jointly form 'Real Essences' shared by all the members of natural kinds.[95] The sense of carving nature at its joints involved in the entire discussion is the non-identity related one.

Brian Ellis, one of the kind theorists presented in chapter 1, follows this rationale. Ellis looks in his expository passages at chemical elements and compounds and compares their micro-structural properties with the superficial properties we can observe.[96] The former properties are to be preferred to the latter, argue Ellis, because the former are intrinsic whereas the latter are relational (the most often cited example of observable, relational property being colour), and also because the former are attended by causal powers, etc. in

---

[95] Of course, as I have already mentioned, apart from the Lockean way, there is a different, semantic vein of argumentation for essences that is often appealed to. In this vein, the term 'gold' for instance, should pick out its reference in all possible worlds in which the reference exists. Since I am primarily looking at metaphysical arguments, this Kripkean vein will not be discussed here in any detail. This was one of the main assumptions of this thesis, as stated in the Introduction to this thesis. I will point out at the end of this subsection how a critique of the Kripkean strategy could directly spin off from the present argumentation. For a fully-fledged criticism of Kripke, see Lowe (2007a), Lowe (2008), Mellor (1977) Salmon (1979) and Salmon (1982). See also Mumford (2005, *esp.* pp. 426, 427) for an argument showing that, on scrutiny, the Kripke/Putnam strategy is not different at all from the Lockean strategy employed by Ellis *et. al.*

[96] See Ellis (2008: 140, 2002: chapter 4). As I have discussed in §1.2, Ellis divides the micro-structural properties in two sorts: the causally efficient properties and the so-called structural properties, namely the properties underlying similarities in shape, spatio-temporal arrangement of various components, etc. How *structural* properties can be part of necessary and sufficient conditions is a problem that cannot be solved by the Lockean strategy. In order to deal with it, Ellis appeals to substantive universals and I shall look at this alternative/complementary strategy in § 2.2.

the sense that they produce the latter.[97] Crucially, Ellis adds, the kinds characterised by such microstructural properties should be non-fuzzy.

'The intrinsic properties and structures of things are *what make them what they are*. They explain how things are disposed to behave, just in virtue of how they are constituted, and what the causal powers of their constituents are. But this is precisely what Lockean real essences are supposed to do. Therefore, the intrinsic properties and structures of things are their Lockean real essences. (Ellis, 2001: 31, italics added)

However, one could point out, the fact that the intrinsic properties and the structure of things explain how things are disposed to behave simply makes those properties determining. Of course, the tag of 'essential' or 'being part of real essences' can be stamped upon them, but nothing in the existence of determining properties as such, justifies taking only those particulars that instantiate non-fuzzy, clear-cut patterns of such properties as natural kind members. Why then should we rule out fuzzy kinds as candidate natural kinds? Ellis seems to be aware of this aspect and continues:

It is important to understand that things could have Lockean real essences even if no natural kind existed...Therefore, for natural kinds to exist, it is not enough that things should have Lockean real essences. Things belonging to different natural kinds must have clearly distinct real essences. There cannot be any borderline cases between the real essences of different natural kinds *because, if there were, the distinctions between kinds would be superficial, as for example, the blue/green distinction is, having no basis in the underlying reality.* For natural kinds to exist, there must be discreteness or discontinuity at the most fundamental level. If objectively different kinds of things exist in nature, the distinctions between them must be there in their intrinsic natures for us to discover.' (Ellis, 2001: 31, italics added)

Ellis notes that we could have 'real essences' possessed by members of certain kinds which are fuzzy (and overlapping) in the sense that the superficial properties and behaviours of such members would be explained by their microstructures. It is just that the microstructures in question would not be exactly similar for all the members of such kinds (and such microstructures would overlap).

However, we see that Ellis dismisses such kinds as un-natural and arbitrary with what appears to be a sleight of hand. Ellis compares the problem of figuring out the membership of individuals in borderline cases (cases in which kinds overlap, or the

[97] See Ellis (2001: 31)

individuals in question do not seem to share a sufficient part of a cluster characterising a fuzzy kind) with the problem of deciding over the mind-dependency of relational properties, like colour supposedly is.[98] This does appear to be a sleight of hand because, if we searched for the way the determining properties agglutinate, the distinctions between the clusters of properties we would find are not 'superficial' and 'with no basis in the underlying reality'. The regions of density one could find in the multi-dimensional space of (determining) properties for medicine and special sciences in general, are not arbitrary and the properties in question are just as real as the determining properties to be found in chemistry. Only if one already accepts the N&S (and non-overlapping) requirement(s) does the possible existence of borderline cases imply that fuzzy (and overlapping) kinds are arbitrary.

Ellis was one of the authors introduced in chapter 1 because his views are indeed representative for a very large number of kind theorists. On the issue of 'essences', Ellis does at least appear to recognise problems that are not even considered by other kind theorists adhering to criterion VIb. The latter are happy to take up the Lockean idiom and claim that nature is surely carved at its joints when we speak about 'essences'. Since this strand of argumentation has been hugely popular in the natural kinds discussions, (so much that the entire debate seemed to be reduced to whether kinds have essences or not), it is worth looking at two more illustrations of it. Take for instance the argumentation employed by Wilkerson in his (1995):

'...something is a member of a natural kind if and only if it has a real essence, that is, an intrinsic property or set of properties that is both necessary and sufficient for its being a member of that kind. The word 'real' crept in, not just as a historical reminder of Locke's distinction between real and nominal essences, but as a philosophical reminder that we are here concerned with essences that are de re, not de dicto, essences that lie in the things themselves and not in our beliefs, thoughts, theories or remarks about them' (Wilkerson, 1995: 109)

What are the 'essential' properties, according to Wilkerson, and how can their existence justify the N&S requirement? Obviously, in order to surpass the chicken and egg dilemma

[98] Parenthetically, what Ellis ignores is that one could take colours as intrinsic properties if one makes the differentiation between the instantiation of a dispositional property and the manifestation of the associated disposition; see Bird (2005) who discusses the respective differentiation.

mentioned in the introductory part of this chapter, tautological definitions of essential properties should not count as satisfactory for grounding this requirement. That is, defining the essential properties as those properties that are possessed by all (and only) the members of a kind should not be considered a convincing justification for why only some of the patterns of determining properties we find in nature, i. e. the non-fuzzy ones, should count as natural kinds.[99] Something else should ground such definitions. What could this additional ingredient be? Wilkerson writes:

'Now, if we are realists [about kinds] we are bound to consider how things are 'in nature' and to wonder whether the kinds distinguished in our theories are natural kinds. To use Plato's familiar analogy, we must wonder whether we are carving at the joints of nature [...] when we identify something as a member of a natural kind, we are in principle able to explain a wide range of its properties, including the apparently superficial properties. The real essence of such a thing not only determines its proper de re classification ....but also directly explains many of its properties...It is precisely because gold has the atomic number 79 that in normal atmospheric conditions is malleable, fusible, soft and heavy' (Wilkerson, 1995: 30, 55)

For Wilkerson then, the essential properties are, evidently, the properties shared by all and only the members of natural kinds. They are also characterized as micro-structural, intrinsic properties that show 'what things are' and (by) explain(ing) the behaviour and properties of kind members. This is insufficient though.

If one puts aside the rhetorical exclamations concerning nature being finally cut at its joints and the things having revealed at last their 'being', there is nothing added to the nature of those properties in Wilkerson's argumentation, aside from the features I previously discussed with regard to the determining properties.

'I suggest that, if we are to produce an interesting account of natural kinds, we should insist that *members of natural kinds, and the corresponding real essences, must lend themselves to scientific investigation*...It is precisely because gold has a certain atomic number that it has certain characteristic properties (its being malleable, fusible, etc.)[...] greetings, nations and banknotes ...do not lend themselves to any sort of precise scientific investigation. *By the same token*, we should insist that natural kind predicates are inductively projectible, whereas other predicates are not. If I know that a lump of stuff is gold, or that the object in front of me is a narcissus, I am in a position to say what it is likely to do next, and what other things of the same kind are likely to do [...] since scientific generalization involves exploring the causal powers of things,

[99] Wilkerson does appear to acknowledge that, without drawing any consequences though; see Wilkerson (1995: 31, 32)

and since causal powers must be constituted or realized by intrinsic properties, the real essences of natural kinds must be intrinsic rather relational properties. '(Wilkerson, 1995: 32-33, italics added)

The majority of kind essentialists employ the same type of arguments. In these arguments, on scrutiny, the 'essential' properties of kinds are nothing more or less than determining properties. Take also Harré's view on kinds, developed originally in his (1974) and recently re-espoused with minor modifications in his (2005). Harré, just like Ellis, focuses on the example of chemical classifications. Harré reminds us in the very beginning what the essential properties for kinds are supposed to be.

"ESSENCES: A REMINDER 1. Properties of kinds a. Proprium: properties found in or displayed by every instance of a kind. b. Essence: properties necessary for a being to be a member of a kind c. Accident: properties found in some instances of a kind. Accidental properties displayed by some individuals in a group cannot be part of the essence of the kind. [...] Essences are sets of properties selected according to various criteria from the propria. In philosophical writings, 'essences' have four main features: i. They are immutable, the essence of a kind cannot change. ii. They are indivisible, a subset of the constituent defining properties of an essences is not an essence. iii. They are necessary, unless a being displays the properties defining the essence it is not a member of the kind. iv. They are infinite, that is an indefinite number of actual beings may realize an essence" (Harré, 2005:10)

Nevertheless, nothing in this generous exposition goes beyond the tautology that essential properties are those properties whose possession is necessary and sufficient for membership. A few passages later, the reader gets a dose of Lockean distinctions and quotes, and is told that the 'essential' properties (of kinds) tell us 'what things are'.

"Though the concept of 'essence' is as old as philosophy, the way that chemists think in terms of essences, though the word may never cross their lips, has its origins in the writings of John Locke (1690) and Robert Boyle (1666). Here is Locke's famous exposition of the distinction between nominal and real essences.. . . its essence, which is nothing but that abstract idea *to which the name is annexed*, so that everything contained in that *idea* is essential to that sort. Whatever we use to distinguish substances into sorts '. . . I call by a peculiar name, the *nominal essence* to distinguish it from the real constitution of substances upon which depends this *nominal essence* and all the properties of this sort, which . . . may be called the real essence' (Locke, 1690: Bk 3, Ch 6, Sect 2)....The real essence of type or kind is the cluster of properties necessary and sufficient for a being to be an instance of a type or a kind" (Harré, 2005: 10-11, italics original)

As in Wilkerson's case, one could ask though - what are the 'essential' properties, beside the fact of their possession being necessary and sufficient for kind membership? Well, it

turns out that they are the intrinsic, structural properties that can be indicated as causative factors for the observable properties and behaviour of their bearers.[100] I conclude that 'Essential' properties, construed in the Lockean tradition, cannot justify the N&S requirement. The Lockean strategy amounts to circular reasoning that defines *de dicto* the divisions of nature corresponding to natural kinds as being non-fuzzy.[101]

As specified in the Introduction to this thesis, I am concerned here with the metaphysical side of the discussion over (natural) kinds. To be sure, one can argue, from a semantic point of view, that the additional 'ingredient' needed in order set aside the *de dicto* charge vis-à-vis the essential properties is the way our *kind* terms behave.[102] As I have already mentioned, there are convincing critiques of this semantics-based strategy intended to justify the N&S requirement but the final part of this sub-section I would like to indicate very briefly my version of the critique.

One can start by defining *kind* terms *simpliciter*, as the terms that refer collectively to particulars and are simply to be distinguished from singular terms (which refer to single particulars). Kind terms, of course, could refer to particulars exhibiting various types of similarities - that is, they could refer to divisions of nature that are non-fuzzy (and non-overlapping) or to fuzzy and overlapping ones. Now, my criticism would be that, *in nuce*, the semantic strategy amounts to categorising *only* the kind terms referring to the former type of divisions as *natural* kind terms. This is why.

The argument employed by Kripke *et. al.* has, in my idiom, roughly the following form:

i) the kind term 'gold' pick up in all possible worlds the *same* division of nature

ii) in the actual world, the respective division of nature fulfils the N&S requirement and the necessary and sufficient conditions are represented by such and such properties

---

[100] Harré, 2005: 12

[101] Several of Locke's commentators (Ayers 1991, Uzgalis 1988, Crane 2003) have underlined this point with regard to his argumentative intentions, as such. Mumford (2005) addresses a critique to Ellis's 'essentialism' that is very close to the argument I present in this section. For kinds to have essences, says Mumford, the conditions of membership should be necessary and sufficient *and* '...there should be something else, something else about the properties characteristic of kinds that makes them 'essential' (Mumford, 2005: 424). Ellis misses the point in his (2005) reply when he re-states his point that basically, the essential properties *are* determining and nothing above (Ellis, 2005[a]: 465-7)

[102] See Soames (2006) for instance

iii) for the respective division of nature to be the *same*, it should be characterised by the N&S requirement,

Hence,

iv) in all possible worlds, the kind term 'gold' picks out the division of nature characterised by such and such properties.

In Kripke's formulation, premise i) says that the kind term 'gold' is rigid and premise ii) says that the reference of 'gold' is discovered in the actual world by *a posteriori* scientific enquiry. Premise iii) says that in virtue of the necessity of identity, terms that are co-extensional in the actual world are co-extensional in all possible worlds and the conclusion iv) states that (metaphysically) necessarily, gold is whatever microstructure is discovered *a posteriori* scientific enquiry.[103] This semantic argument, if valid,[104] could be used in order to infer something about the metaphysical category of natural kinds only if the following corollary was added:

a) 'gold' is a *natural* kind term

b) 'gold' refers to a division of nature that necessarily fulfils the N&S requirement;

hence

c) necessarily, *natural* kinds fulfill the N&S requirement.

However, to hold this corollary is to categorise *only* the kind terms that refer to the non-fuzzy type of divisions of nature as *natural* kind terms.[105] And this is just a linguistic formulation of the rationale which, I have shown, is underpinning the metaphysical arguments of the theorists adopting the Lockean strategy.

---

[103] See Kripke (1980) and Soames (2006)

[104] and there are serious reasons to think that it is not, *Cf.* Stewart (1990) for instance, but this aspect need not concern us here

[105] To stress, this critique represents merely a re-formulation of the general critique addressed to Kripke by various authors; see for instance Lowe's (2007) and (2008) discussion.

§2.1.2 *ESSENTIAL PROPERTIES AND IDENTITY*

I have argued in the previous sub-section that 'being essential' is a simple tag employed in connection with certain properties, in order to make the N&S requirement acceptable. I have shown that the supposed justification for this requirement is circular. Natural kinds, it is claimed, could only be those that possess essential properties, where essential properties are those part of necessary and sufficient conditions. We have here two definitions that sustain each other in what reveals itself as a stipulation about the metaphysical category of natural kinds – a definition about those kinds that are natural and about the type of properties characterising the natural kinds. The tag of 'essential' attached to the determining properties is unjustifiably linked with a label attached to the category of natural kinds, as it were.

To be sure, the informal way of speaking about 'tags' and 'labels' in this context might not be the most happy choice. A more adequate analysis of the problem is to say that theorists like Ellis, Wilkerson and Harré - who work in the Lockean tradition and, I should add, employ (actively) a certain sense of carving nature at its joints – make use of an equivocation between different meanings of 'the essential'.

I have differentiated in chapter 1 between three main senses of the essential that should be neatly distinguished – i) as explanatory powerful (on the epistemic side) or causally efficient (on the metaphysical side) ii) as part of N&S conditions and iii) as contributing to the identity of kind members.[106] The equivocation at issue involves in its argumentative ramifications all three senses of the essential. So far, I have described the back and forth between i) and ii), which amounts to categorising properties that are (minimally and) adequately picked out by i) as falling (also) under the heading of ii) and hence leads to a *de dicto* reasoning about the metaphysical category of natural kinds. Let there be a kind K characterised by the pattern of determining properties [p_1, p_2, ...p_n] and an individual *a*

[106] Just to make matters worse, Ellis refers sometimes to the N&S requirement using roughly the phrase 'identity of kinds' (See Ellis 2001: 19, 52, 75, 246). However, for keeping decently clear a terminology that's already entangled (in the literature and in my own discussion as well) I shall continue to refer to criterion VI_b as the N&S requirement. Ellis's idiom for the N&S requirement is worth bearing in mind though because it is I believe connected with a certain tacit use of the identity-sense of carving nature at its joints, a use that I shall return to a bit later.

that instantiates all of [p_1, p_2, ...p_n]. To employ the inference - if *a* lost, say, the property p_3, then it would not be any more a member of the kind K, were K to be natural - amounts to a *de dicto* decision (based on the semantics of the ii) sense of the essential) about the nature of the metaphysical category of natural kinds.[107]

However, the equivocation at work in the arguments of Ellis *et. al.* involves the iii) sense as well, mainly because the N&S requirement could be justified only if the identity of *a* depends on all of the [p_1, p_2, ...p_n] properties. If it did depend, then *a* could not lose p_3 and remain a member of the kind K, for the simple reason that after the change, we would have instead of *a* another individual. Hence, to explicitly invoke the ii) sense of the essential and claim that kinds should have necessary and sufficient conditions of membership is also to equivocate between ii) and iii). In fact, this last equivocation boils down to appealing tacitly to the identity sense of carving nature at its joints and I will come back to this aspect shortly. Prior to that, I should add a few remarks about criterion VII.

I have said in the beginning of this chapter that I will treat criterion VIb and VII (N&S and non-overlapping requirements) together. Most of the time, criterion VII is advanced complementary to criterion VIb in that it is assumed that if something fulfils criterion VII, then it also fulfils VIb. In my view, they are (also) related in the sense that their only valid justification would have to appeal to the identity of their members.

There are numerous criticisms that can be addressed to criterion VII. Amongst them, one of the simplest and most effective merely points out that the exact sciences are rife with criss-crossing classifications - think for instance of how elements in chemistry could be classified into metals/non-metals and solid/liquid/gaseous. [108] However, we can also advance the same type of argumentation I have employed vis-à-vis criterion VIb, with respect to the presence of a certain vicious circle.

---

[107]This is somewhat ironical, if we recall that a basic ingredient of the Lockean strategy is to contrast the 'real essences' with the 'nominal' essences and emphasize that the latter are merely *de dicto* definitions that could at best be useful for recognizing natural kinds; see Wilkerson (1995) for instance.

[108]See Khalidi (1998: 40,41,50). Khalidi also notes in his (1993) discussion of kinds as carving nature that criteria VIb and VII are connected.

Let there be two kinds K and K* characterised by the patterns $[p_1,p_2,....p_n]$ and $[p_2, p_3....p_n, p_{n+1}]$ respectively, which overlapped (without being in a hierarchy), in the sense that an individual a instantiating $[p_1,p_2,....p_n, p_{n+1}]$ was a member of both. To argue that at least one of them could not be natural, is to stipulate something about the metaphysical category of natural kinds, insofar as the properties involved are just determining properties.[109] The non-overlapping requirement would be justified, again, only if the identity of (natural) kind members depended upon their membership. Suppose that a lost the property p_{n+1} in a process characterised by spatio-temporal continuity, continuity of causal history, etc.[110] From the point of view of its membership in K, the individual resulting from the process should be (numerically) identical with *a*. From the point of view of its membership in K*, the individual in question should be different from *a*, since the kind membership (one of the necessary conditions for identity preservation) was lost.[111] We would thus have a violation of the *tertium non datur* principle.

What the non-overlapping and N&S requirements can at best achieve is to avoid certain intuitive 'paradoxes' of criss-crossing and fuzziness, by precluding the 'essences' of kinds from 'mixing up' or 'being chopped'. These 'paradoxes' are associated with a certain construal of kinds as 'entities'[112] and are also made possible by a tacit appeal to a certain set of assumptions. Here again is Ellis, explaining why chlorine is a natural kind, given that it fulfils the two requirements.

---

[109]See for instance Ellis (2005b:376, 2005a:463,464, 466) in which Ellis just argues that natural kinds should be mind-dependent and that 'on inspection, it is clear that the ways in which reality is divided are not mutually independent, but have a natural hierarchical structure' (Ellis, 2005a: 466).

[110]One could mention here any other conditions, beside natural kind membership (our present hypothesis), upon which various authors think that the numerical identity of individuals might depend. Ellis (2001: 239) mentions only causal and temporal history for instance. Lowe introduces beside spatio-temporal continuity certain criteria specific for three categories (material objects, artefacts and living beings) like the preservation of matter, for material, inanimate objects; see Lowe (2001: 55, 56,174 -178). More on Lowe's criteria of identity will be discussed in § 2.3

[111] See also Mackie (1994: 323). I have framed this example to suit the views of someone who adhered both to criteria VIb and VII because usually the theorists who adhere to one also stick to the other. But the example could be adapted also for the case in which VII was upheld and VIb rejected, that is, the case in which K and K* were cluster kinds. In this alternative scenario, we could imagine *a* to lose not only p_{n+1} but a sufficient part of the pattern $[p_1,p_2,....p_n, p_{n+1}]$ such that it would not be any longer a member of K* but it would still be a member of K.

[112]The assumption that kinds are some sort of entities is mentioned in Bird and Tobin (2008), in direct connection with (what I have called) requirements VII and VI_b. For an earlier reference, see for instance Donnellan's discussion of Putnam's essentialism in Donnellan (1973)

'What is true of the chemical kinds is not true of biological species. The existing species of animals and plants are clusters of morphologically similar organisms whose similarities are due to their genetically similar constitutions. Our species concepts are therefore generic cluster concepts. They are not, however, generic kinds that are categorically distinct from one another, as the generic chemical kinds are. The species "elephant" has a number of sub-species, which are sub-clusters within the elephant cluster. These sub-species are distinct enough to be reliably distinguished morphologically, and sufficiently different genetically to be said to be different kinds of animals. However, if we broadened our vision to include all of the ancestors of the current elephants in the world, we should find, I think, that the morphological clusters, and the genetic clusters that explain them, would shift about as we go back in time, and would eventually overlap. Therefore, neither the generic species nor any sub-species of elephant is a natural kind in the same sense as the generic and specific chemical kinds are. Chlorine, for example, is a generic chemical kind, the species of which include the various isotopes of chlorine. But there is no species of chlorine existing now or at any other time that could possibly be a species of any element other than chlorine. Chlorine, the generic kind, has a fixed nature, and each species of chlorine has its own fixed nature. (Ellis 2008: 140-1)

Why do the 'mixing and chopping' of 'essences' appear paradoxical?[113] Because, I further wish to claim, the identity of kind members is tacitly appealed to in framing such scenarios. Take the classical example of the cyclist and the mathematician paradox formulated by Quine,[114] which perfectly parallels the intuitive image that Ellis appeals to. If we accept, says Quine, that mathematicians are necessarily human and not necessarily two legged, and cyclists are necessarily two legged and not necessarily rational, it would follow that a person who is both cyclist and mathematician is and is not necessarily rational, and is and is not necessarily biped. Quine regards such 'criss-crossing' situations as paradoxical (which, in our context, parallel the paradox of 'mixing' of essences for kinds which fail the non-overlapping requirement) and employs them in order to reject the notion of 'essence' (and the associated notion of necessity).

Ellis appeals, as we have seen, to similar examples, which are also regarded as paradoxical (and in order to see this more clearly, we need only think of 'mathematicians' or 'cyclists' as candidate kinds). However, Ellis prefers to retain the notion of essence, and dissolves

[113] For clarification, in this particular context of intuitive 'paradoxes' arising out of violations of criteria VI_b and VII, I employ the term 'essence' with regard to the construal of essential properties are causally efficacious properties – the determining properties. On the other hand, what I ask is how could authors like Ellis *et. al.* regard the 'mixing' of putative essences for candidate kinds (violations of criterion VII) and the 'chopping' of them (violations of criterion VI_b), as paradoxical without simply reciting the respective criteria. There is an intuitive source for an apparent paradox, but what is it?

[114] Quine (1960: 199)

the paradoxes by holding that essences cannot be mixed or chopped up, that is, by adhering to criteria VIb and VII. Now, we get to the heart of the matter when we observe that Quine's envisaged situation is paradoxical indeed only if modal statements such as 'mathematicians are necessarily rational' for instance, are given a *de re* reading.[115] That is to say, only if the identity of persons who are mathematicians depends on them being rational and the identity of the persons who are cyclists depends of them being biped, a person who was both mathematician and cyclist and went mad or lost one leg would remain, and would not remain identical with itself. On a *de dicto* reading of the respective statement(s), no paradox arises. Any human being can be both a cyclist and a mathematician, just as, we could add, any particular can be member of two overlapping kinds.[116]

Quine would reject this analysis of his paradox, given his adherence to referential opacity in modal contexts.[117] Ellis could not invoke the same reason, given that modal contexts are at the centre of his overall argumentation. Ellis would have two options, in order to ground his assumption that the chopping or mixing of essences is paradoxical. The former would be to explicitly stick to the *de re* reading of the (first order quantified) statements espousing the conditions of membership for kinds, and make the identity of kinds members dependent upon their memberships. The second would be to maintain the 'de dicto' reading, as it were - that is, quantify the respective statements in a second-order way. The latter move would transmute the violations of the *tertium non datur* principle from the case of individual kind members with their numerical identity, to the case of kinds and their identity (where kinds would be substantive universals or

[115] See for instance Oderberg (2001: 28-34, esp. p. 32), Marcus (1993: 227).
[116] i.e. can possess the 'essential-for-the identity-of-kinds' properties of two different kinds. Recall that Ellis calls sometimes the 'essential' properties associated with the N&S and non-overlapping requirements as 'essential for the identity of kinds' or 'essential for kinds'. As I said, this usage is important for me when it comes to discussing the tacit use of a certain sense of carving nature at its joints.
[117] See Hylton (2007: 353, 354), Teller (1975:233, 234). We shall return to the notion of relative identity in §3.6 As I said, until the respective section, the notion of identity used in my text is the absolute identity, unless otherwise stated.

whatnot).[118]

To stick to the latter option would indeed mean to acknowledge that the conditions of (natural) kind membership do not have to be fulfilled by all and only the members of a kind (at least not a result a of reduction ad absurdum argument, concerning the above paradoxes). To stick to the former option would be to introduce the identity of kind members into the equation. Ellis does neither. That is, neither the identity of kind members is taken into account, nor the second order quantification is adopted, yet the N&S and non-overlapping requirements, as well as the talk about the properties 'essential-for-the-identity-of-kinds' are maintained.

What this means, I hold, is that the identity of kind members is tacitly appealed to in framing the putative paradoxes of criss-crossing or fuzziness. Of course, the examples regarded by Ellis as paradoxical have a certain intuitive appeal, just like regarding kinds as 'entities', which brings us back to the related point concerning the assumption that natural kinds carve nature at its joints.

When discussing earlier the case of Ellis *et. al.* vis-à-vis 'essences' and the disguised determining properties, I have insisted on their use of the expression (and their claim that) 'what things are' is revealed by such properties (also constituting for them necessary and sufficient conditions for membership). Now, notice that to construe in the above-mentioned identity-related way the relation between kind members and the kind properties they instantiate is to adhere to the identity sense of carving nature at its joints. Wilkerson, Ellis and Harré, however - as almost all the theorists holding that natural kinds have necessary and sufficient conditions of membership fulfilled by all and only their members[119] - explicitly adopt the non-identity related sense of carving nature at its

[118] Indeed, a second order quantification concerning kinds *simpliciter* would presumably range over substantive universals. I shall move to substantive universals in the next sub-section where we shall see that they cannot ground requirements VII and VIb either. Of course, such quantification would be as it were 'de dicto', just in the sense in which the entities quantified over would not be particulars / particular substances.

[119] One exception is Wiggins (2001). Some problems with the type of argumentation that Wiggins and others employ will be discussed in §3.6

joints.[120]

I conclude that their argumentation vis-à-vis essences uses tacitly the other sense of carving nature at its joints. Only if this (identity-related) sense was adopted, the N&S and non-overlapping requirements (criteria VIb and VII) could be justifiably employed in the Lockean strategy to pick out kinds and the overlapping and fuzziness of kinds would indeed produce paradoxes.[121]

In the next section, I will look at another possible justification of requirements VIb and VII, revolving around substantive universals. Substantive universals could be invoked as the necessary *'de re'* ingredient needed in order to overcome the *de dicto* reasoning of the Lockean strategy. Notably, Jonathan Lowe – the philosopher who reintroduced substantive universals into the kinds discussion - does not hold that on the level of particulars, the divisions of nature standing for the entities instantiating these substantive universals should be non-fuzzy and well-delineated.[122] However, Brian Ellis has appealed to such substantive universals as justifications for requirements VIb and VII, especially in order to solve the problems that structural properties pose for his account of metaphysically necessary laws.[123]

Could substantive universals fare any better than the Lockean strategy in justifying the existence of necessary and sufficient conditions of membership for any natural kind? I will argue that they could not, on three related scores. First, I will show - vis-à-vis the very introduction of substantive universals into our metaphysics of (natural) kinds - that Ellis's specific argumentation suffers from the same shortcomings as the general Lockean strategy, in that the assumption of natural kinds as standing for non-fuzzy and non-overlapping divisions of nature is simply taken for granted. Second, by comparing Ellis's

---

[120] Both Wilkerson and Ellis reject that the identity of kind members has to do with their kind membership. Harré does not even discuss the issue of identity; see Wilkerson (1995) Ellis (1999: 67,68) (2001:239), Harré (2005)

[121] This is not meant to deny that *some natural kinds do actually have necessary and sufficient conditions of membership* (and possibly have these conditions of membership in all possible worlds). My point is that not *all* natural kinds need to have necessary and sufficient conditions of membership, on the non-identity sense of carving nature; that is to say, the N&S requirement is not justified for the natural kind *category* if the identity sense of carving nature is not adopted.

[122] See Lowe (1989: 153-154), Lowe (2006: 13-17)

[123] Ellis (2001), (2002), (2005), (2008)

views with Lowe's, I will show that even if we accept the existence of substantive universals, specific details of these universals' metaphysics make the N&S requirement very problematic. More precisely, I will argue that Ellis fails to explain, in *his* framework of kinds, what Lowe's relation of 'characterisation' between substantive universals and non-substantive universals amounts to. Third, I will show then that Ellis does not solve the difficulty that the perfect sameness of categorical structures for kind members poses for his account of metaphysically necessary laws.

§2.2 SUBSTANTIVE UNIVERSALS AND REQUIREMENTS VIB AND VII

In the metaphysics of natural kinds re-introduced into the philosophical debate by Jonathan Lowe, natural kind membership is not primarily determined by the possession by particulars of certain properties but by their instantiation of substantive universals. Only secondarily (in an ontological sense) do kind members come to possess kind properties, due to a certain relation of *characterisation* holding between substantive and non-substantive universals. To use one of Lowe's expository examples, a particular apple is a member of its kind primarily because it instantiates the substantive universal apple. The particular in question is red, round, solid, etc. as a kind member just because the non-substantive universals redness, roundness, solidity, etc. 'characterise' the substantive universal apple.

As mentioned in §1.2, Lowe does not hold that members of a natural kind, viz. particulars instantiating a certain substantive universal, should also instantiate a complete set of properties whose possession is necessary and sufficient for membership. Lowe's approach to these kind properties is situated at the level of possibility. In the particulars that are members of a natural kind, a certain set of properties *could* be instantiated. The more such properties are instantiated, the more kind members come to resemble the substantive universal (which is perfectly 'characterised' by the substantive universals specific for the kind in question).[124] We might have particular apples that are not perfectly round, particular ravens that are not black, etc. even if the substantive universal apple is

[124] Lowe (2006)

characterised by all the non-substantive universals corresponding to the qualities we associate with apples.

Brian Ellis's recent work on natural kinds acknowledges the existence of substantive universals and requires at the same time that the divisions of nature corresponding to natural kinds on the level of particulars be non-fuzzy. That is, Ellis wants the N&S requirement to hold for natural kinds. Of course, this is not the only difference between Ellis and Lowe's views. Among others, Ellis also differs from Lowe in the way of construing the dispositional/categorical distinction. Broadly speaking, for Lowe the distinction is grounded in our mode of predication (such that all the properties actually instantiated by kind members are categorical, as it were). For Ellis, the properties that kind members actually share are split into causally efficient properties *and* structural properties - the latter underlying similarities in shape, spatio-temporal arrangements of various components, etc. Just as in Lowe's case though, these properties, found on the level of particulars, represent tropes of non-substantive universals. Their possession is due to the instantiation of a substantive universal and, presumably, to a certain relation between the respective substantive universal and the non-substantive universal (a relation whose nature, as we will see, Ellis hardly discusses).

Interestingly, Ellis seems to have adopted substantive universals mostly in order to explain the presence of 'identical' structures on the level of kind members. That is, the main use of adopting them is to underpin and justify the N&S requirement on the level of *structural* properties. For instance, a substantive universal should underpin/explain the fact that all the members of the methane kind present the same shape of the CH_4 molecule. In Ellis's terminology, *infima species* of substantive universals should have 'freestanding' tropes, acting like a sort of spatio-temporal 'skeleton' for the tropes of the power non-substantive universals that are characteristic for each natural kind.[125] With respect to the causally efficient (determining) properties, the role of metaphysical underpinning for the perfect similarities between kind members is played by 'essences' and 'essential' properties, along the path of reasoning I have described as the Lockean

[125]See Ellis (2001: 73-75, 67-69)

strategy.[126]

We should note that there are certain questions to be asked regarding the very plausibility of accepting the existence of substantive universals. In Lowe's scheme for instance, such universals should play the role of truth-makers for scientific laws, constitute grounds for co-instantiation inductive inferences about kind members, respond to the threat of Heraclitanism or Spinozism about particulars etc.[127] It has been argued, however, that we have more metaphysically parsimonious ways of dealing with these issues.[128] For Ellis, as I have already remarked, the main reason for introducing substantive universals is the need to justify the N&S requirement for natural kinds.[129] Two related queries can then be raised about his position. First, why should natural kinds be substantive universals, especially in the context in which natural kinds are supposed to fulfil the N&S requirement? And second, how could substantive universals warrant the N&S requirement at the level of particulars?

The former query concerns the very introduction of substantive universals into one's metaphysics. As we will see, it brings to light the same type of *a priori* reasoning about natural kinds I have emphasized vis-à-vis Lockean 'essences' and 'essential' properties. The latter is related to and yet independent from the former in that one could accept in principle the existence of substantive universals and yet reject the fact that something like the N&S requirement follows from their existence. I will look in turn at these two queries in §2.2.1 and §2.2.2 and then will analyse in §2.2.3 the particular case of structures and structural properties vis-à-vis this requirement.

---

[126]The complementarity between the Lockean strategy and the substantive universals strategy can only be observed in his (2001) treatise. In Ellis's (2002) book substantive universals are hardly mentioned.

[127] Lowe (2006: 135)

[128] For one, the truth-makers of laws could simply be taken to be properties (Bird 2007). For another, the co-instantiation aspect can be explained without inflating our ontology, by appealing to properties and causal mechanisms, either internal or external, which produce the stability or unity of the determining properties at the level of kind members (Boyd 1991, Bird 2007). Also, Heraclitanism or Spinozism can be refuted by interpreting the individuation of particulars in terms of spatio-temporal continuity, sameness of causal history, etc and/or by viewing kinds as participating in individuation without kinds being substantive universals; see Wiggins (2001), Wilson (1999), Hirsch(1982). More on Heraclitanism and Spinozism will be discussed in §2.3

[129] Ellis also wants kinds to be the truth makers of laws but for various reasons (having ultimately to do with his distinction between categorical and dispositional), his kinds are not really the truth-makers of his laws – powers are; see Botterill (2005)

§ 2.2.1 WHY HAVE SUBSTANTIVE UNIVERSALS IN OUR ONTOLOGY OF NATURAL KINDS?

This is how Ellis sets the stage for the introduction of substantive universals:

'*The most salient general feature of reality I seek to explain is its natural kind structure.* There are hierarchies of objects of increasingly complex kinds....The more complex structures are compounded of simpler structures, the more complex properties, of simpler ones. Within each of these hierarchies, we may detect various *samenesses* (*exact* resemblances)....These identities *can best be explained* by *assuming* that there are universals *that have particular objects....or structures as their instances*....These similarities can best be explained by assuming that there are hierarchies of generic universals, the infimic species of which are all specific universals with *identical* instances. An ontology in which the *only* property universals that are recognized as fundamental are non-structural ones does not seem to be up to the task of accounting for the complex structures that are to be found in nature' (Ellis, 2001: 67-68, italics added).

We see that, for Ellis, the *explanandum* is the fact that *natural* kinds and *natural* kind members satisfy the N&S (and hierarchy) requirement(s) and the *explananans* is the introduction of substantive universals (beside property universals). His 'methodology' thus raises doubts from the start. I am not referring here to his (general) adoption of an Inference to the Best Explanation strategy for solving metaphysical disputes,[130] but simply to the way in which Ellis frames the very phenomena to be explained by the introduction of substantive universals. What raises doubts, more precisely, is that we seem to have here precisely the *a priori* reasoning present in the case of Lockean 'essences'. Ellis thinks we should seek to explain the perfect similarities between (certain) natural kind members by postulating substantive universals. Nevertheless, should we not explain in the first instance why *only* the divisions of nature that present these perfect similarities between the involved individuals are to be acknowledged as natural kinds?

In different terms, the dilemma raised by Ellis's use of Inference to the Best Explanation goes as follows. Are the divisions in nature corresponding to natural kinds non-fuzzy (and non-overlapping) because natural kind members are said to instantiate substantive universals? Or should natural kind members instantiate substantive universals because the divisions in nature corresponding to natural kinds are non-fuzzy (and non-

[130] Mumford (2005) offers a critique of Ellis's conclusions by focusing on his Inference to the Best Explanation strategy, and taking into account the entire metaphysical picture that Ellis draws in his 2001 treatise.

overlapping)? Ellis obviously thinks that the latter question is worth asking and responds affirmatively. The non-fuzzy (and non-overlapping) structure of natural kinds represents, says Ellis, 'a most salient feature of reality'. The former question is overlooked. This means that Ellis sets aside the possibility of construing in any other way the divisions in nature corresponding to natural kinds. If one considered this possibility, then the fact that the natural kinds satisfy the N&S requirement should at least become less salient and accordingly, the need to introduce substantive universals in our metaphysics of natural kinds would lose much of its urgency.

The same assumptions are present in the details of Ellis's (otherwise remarkably rich and carefully worked out) metaphysical foundation for natural sciences. Ellis writes for instance:

'The instances of each infimic species of natural kinds [i.e. substantive universals] in the category of substances must be all essentially the same. *For if they were not, then the species would have subspecies.* They are therefore like classical universals (which have no sub-species) although, as we will see, they are not property universals either' (Ellis, 2001: 70)

Of course, *if* natural kinds *should* fulfil the N&S requirement, then every time the members of a putative kind are not 'essentially the same', we must suspect that the kind in question groups together, on arbitrary grounds, members of different other *natural* kinds. Hence, we must suspect that it does not point to the instantiation of a specific substantive universal by the particulars in question. Nevertheless, in the above passage, as in all the other places in which the N&S requirement is discussed vis-à-vis substantive universals,[131] Ellis seems to claim that natural kinds as substantive universals should fulfil the N&S requirement *because*, if they did not, they would not be natural.

We could easily argue in favour of rejecting the introduction of substantive universals into the metaphysics of kinds, at least if the reasons for introducing them are Ellis's. The problems faced by his argumentation are not reduced, however, to the general problems of the Lockean strategy. Vindicating the N&S requirement also proves very problematic if we consider some details of the 'functional' ontology of substantive universals.

[131] e.g. Ellis (2001: 52-54)

§ 2.2.2 How could substantive universals justify the N&S requirement?

Suppose we accept for the sake of the argument that substantive universals need to be introduced into our metaphysics of kinds. After all, there might be other reasons I have not reckoned with that could compel us to accept their existence. But how precisely could the substantive universals justify the exact similarities between members of natural kinds (i.e. individuals instantiating the same substantive universal)? What I want to indicate is that even if we accept their existence, there is no explanation on Ellis's behalf for why the kind members should instantiate *all* of the properties characteristic of a natural kind.

It is important here to recall that Lowe, the main proponent of substantive universals in contemporary metaphysics, does not adopt the N&S requirement. In Lowe's scheme we have relations of 'characterization' between *all* the non-substantive universals of a kind, on the one hand, and the substantive universal representing that kind, on the other. These relations cannot be 'translated' at the level of particulars though. The latter *can* instantiate *all* of the properties specific for a natural kind, but they do not have to.[132]

Ellis takes on the main insight of Lowe's metaphysics of kinds and we have seen how Ellis's argues in favour of accepting substantive universals into our ontology, as the best explanation we have for the fulfilment of the N&S requirement by kind members. Nonetheless, Ellis does not examine at all how any substantive universal could possibly ground from a metaphysical point of view the presence of perfect similarities between the particulars instantiating it. All we have on his behalf are a series of declarative statements, of the form:

> '[the perfect similarities between kind members] can best be explained by assuming that there are hierarchies of generic universals, the infimic species of which are all specific universals with *identical* instances.' (Ellis 2001: 67-68, italics added)

Surely, the most direct way of searching for such a justification would be to look at the

[132] For Lowe, this is in fact one of the advantages of introducing substantive universals; see Lowe 1989: 153-154.

relation of 'characterisation' (to choose the name Lowe employs, for simplicity's sake) holding between substantive universals and the non-substantive ones. Ellis does not discuss at any point this relation in his 2001 treatise (or, elsewhere for that matter). What could possibly be the relation of 'characterization' between, say, a substantive universal NK and three non-substantive universals P_1, P_2, and P_3, such that any individual instantiating NK *should* also *co*-instantiate the corresponding properties p_1, p_2 and p_3, in the sense specified by the N&S requirement?

It should be emphasized at this point that the explanation for why certain determining properties are agglutinated in perfect similarities between members of kinds (thus *describable* by necessary and sufficient conditions of membership) should not appeal to causation. That is, the relation of 'characterization' should explain the perfect *co*-instantiation of (determining) properties in terms that do not appeal to the Boyd-like model of divisions of nature resulting from the clustering of properties due to the causal structure of the world. Such a causal structure might determine a perfect clustering of properties tantamount to perfect similarities holding between particulars, but could not explain how the N&S requirement holds for certain kinds *specifically* due to their instantiating a substantive universal.

One possibility would be to appeal to a part-whole relation. In this case though, it would be very difficult to maintain the ontological status of substantive universals, as entities standing one level higher in the hierarchy of universals, over and above the non-substantive characterising kinds.[133] This move would make NK a conjunctive, non-substantive universal, having as parts the universals P_1, P_2, and P_3 and its substantial role status would be lost. Ellis would certainly reject this possibility since he rejects conjunctive properties.[134]

Another possibility would be some sort of relation of necessitation.[135] Nevertheless, such relations of necessitation between universals have a bad name. First, it is because they

[133] *Cf.* Armstrong (1997)
[134] Ellis (2001: 68, 91, 92, 96-7)
[135] in Armstrong's (1985) sense.

seem *ad hoc* in general.[136] In our case, to say that such a relation holds between a substantive universal NK and the non-substantive universals P_1, P_2, and P_3 is no more than to posit that kinds should have necessary and sufficient conditions of membership. Second, because it is unclear how such a relation between *universals* is 'transmitted' at the level of *particulars*.[137] It might be that this 'transmission' aspect gains some plausibility if the correspondent of the nomic necessitation between universals at the level of particulars is taken to be the immanent necessity underlying the successions of properties.[138] But in our case, we are not talking about causation. The relation between the properties p_1, p_2 and p_3 is one of *co*-instantiation and this is precisely what needs to be explained by the relation between NK and P_1, P_2, and P_3.

Perhaps one could argue – in analogy to a well-known rationale concerning the relation between non-substantive universals and causal powers[139] – that the identity of a substantive universal depends on its 'association with' a certain set of non-substantive universals. The identity should be dependent *the sense in which* the instantiation of a substantive universal NK by a particular *a* should bring about the co-instantiation by *a* of a set of properties p_1, p_2 and p_3 (corresponding to non-substantive universals P_1, P_2, and P_3).[140]

But the analogy could not work, I think. In the case of the association between powers and properties, the argument is backed up by the Eleatic principle (namely the principle that no entity that does not make a causal contribution should be accepted in one's ontology).[141] It is also backed up by the serious semantic and epistemological problems of the rival conception of causation and property individuation.[142] However, in our case,

---

[136] See Mumford (2004: 93) Lange (2000: 8)

[137] This is the famous 'problem of inference' to the level of particulars; see for discussion Handfied (2009: 294), Cartwright and Alexandrova (2007: 796) Mumford (2004: 93)

[138] *Cf.* Armstrong (1997)

[139] Shoemaker (1998); see also Bird (2007:71) Armstrong (1985:161)

[140] One could argue that the identity of a substantive universal depends on its characterizing non-substantive universals, without this entailing anything about what properties kind members should instantiate. Indeed, this is what Lowe intimates (Lowe, 2006: 155, 169, 170, 173). This strand of argumentation would not be of much help for Ellis, since the N&S requirement concerns the level of *particulars* and their properties.

[141] See Colyvan (1998) for a (critical) discussion.

[142] Problems issuing from the possibility that the properties could associate different causal powers see Bird, 2005)

again, between p_1, p_2 and p_3 (and whatever other property stands for the instantiation by *a* of NK) we do not have causation but *co*-instantiation and no principle analogous to the Eleatic one is available vis-à-vis the individuation of substantive universals.

Now, in addition to using, on a general level, Inferences to the Best Explanation in ways that interpret in advance the category of natural kinds, Ellis also employs a specific argumentation concerned with the *structural* properties of kind members. This argumentation is worth analysing because it constitutes a very good illustration of the tribulations faced by the N&S requirement when the nature of the relation of 'characterisation' is not clarified.

§2.2.3 STRUCTURAL PROPERTIES AND THE N&S REQUIREMENT

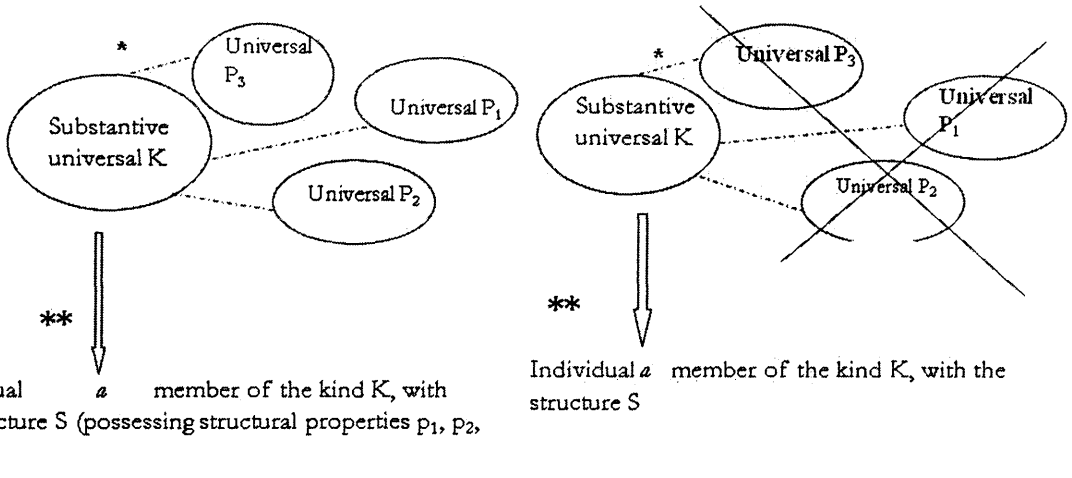
My hypothesis is that in the overall argumentation that Ellis employs in his (2001), the relation of 'characterisation' from the level of universals is simply taken as redundant and unworthy of further consideration. For Ellis, the standards imposed by the N&S requirement on the co-instantiation of properties, taken as causal *powers*, are simply to be justified by the Lockean strategy. As for the perfect co-instantiation of *structural* properties, Ellis appears to think that it is sufficient to invoke the relation of *instantiation* between particulars and substantive universals *simpliciter*, without any need to examine the relation between the (structural) properties and the substantive universals via the non-substantive universals, as in Lowe's case.[143] What we can see in Ellis's argumentation, when comparing it to Lowe's, is a certain change of perspective, as it were, which can be indicated in a simple, heuristic notation as follows.

For the case of a structure S possessed by a particular *a*, what we observe when Lowe and Ellis's views are compared is a change of focus *from* the relation of characterisation of the substantive universal K by, say, non-substantive universals P_1, P_2 and P_3, *to* the relation of

[143] Lowe argues that structural properties such as shape are simply non-substantive universals (just like mass or charge), which characterize the kind, substantive universals. Lowe does not call the structural properties *categorical*, because he employs a different construal of the categorical-dispositional distinction. See Lowe (2006: 17, 124-127)

instantiation by *a* of the substantive universal K, *simpliciter*. In other words, instead of the structure S being possessed by *a* due to its instantiation of K, due to the relation of characterisation between P_1 P_2, P_3 and K and due to the structural properties p_1, p_2, p_3 possessed by *a* representing tropes of P_1 P_2 and P_3, Ellis simply invokes *a*'s instantiation of K.

***FIGURE* 4 The switch of focus from the relation of characterisation of the substantive universal K by non-substantive universals P1, P2 and P3 (as it appears in Lowe), to the relation of instantiation by *a* of the substantive universal K, *simpliciter* (as it appears in Ellis)**



The change of focus from the 'characterisation' relation to the relation of instantiation of substantive universals *simpliciter* is realised by way of a series of interrelated arguments.

Ellis contends that structural universals cannot be reduced to non-structural universals,[144] that the instantiation of structural universals presupposes its bearers possessing the *same* structures,[145] that the instantiation of non-substantive universals does not presuppose their bearers being the *same*,[146] and hence that the properties possessed by kind members *generaliter* (i.e. their resemblances) cannot be satisfactorily grounded in non-substantive universals only.[147]

Substantive universals are thus (re-)introduced by way of 'substantivising' structures,[148] as it were, and two remarks should obviously be made. On the one hand, the possibility of reducing structures to structural non-substantive universals is not considered.[149] On the other, again, the main reason for taking substantive universals as grounds for the structural resemblances between (natural) kind members is that the N&S requirement

---

[144] See Ellis (2001: 68). *Contra* Bigelow and Pargetter (1990), Ellis argues that the universals standing for the atoms of carbon and hydrogen, cannot account for the shape of the methane molecule, because 'universals cannot be spatially arranged'.

[145] 'As soon as we speak about spatial arrangements of objects, we are already invoking the language of structural universals, since different groups of objects can presumably be arranged in exactly the same way' (Ellis 2001: 68-69)

[146] "The instances of property and relation kinds are tropes; those of substantive universals are objects or substances....The instances of any specific quantitative property are *property instances*; they are not the objects in which these instantiations may occur, because the objects in which such a specific universal is instantiated need not all be the same...Two grams of water is not identical to two grams of alcohol....Therefore the instantiations of the classical universal in these two cases are not the samples, but the masses of the samples' (Ellis 2001: 70, 71, italics original)

[147] '...one might think of the property of being a hydrogen atom as a simple conjunctive universal (which is the simplest kind of structural universal). A thesis of this kind is certainly plausible for the fundamental constituents of matter, which presumably have no structure. But hydrogen atoms and the atoms and molecules of more complex substances all have structures, some of them quite complex. Yet their structures are no less essential to their identities as kinds' (Ellis, 2001: 74, 75) One wonders though, if 'universals are not spatially locatable' (p.75), how come substantive universals can fulfill for structures possessed by particulars the role that structural, non-substantive universals cannot. Ellis acknowledges in the final part of the respective section dedicated to substantive universals and structures that 'tropes have the advantage of being spatially located, and relatable by primitive relations'. Campbell's (1990) attempt to reduce structures to tropes is dismissed for the main reason that 'it does nothing to explain the broad facts about the world alluded in the Basic Structural Hypothesis' (Ellis 2001: 76), where the Basic Structural Hypothesis is the thesis that the world possesses the structure of natural kinds Ellis thinks it possesses (see Ellis 2001: 22). No mention of Lowe's work is made in the entire treatise.

[148] See Ellis (2001: 73,75 *et passim.*) for the continuous vacillation between designating structures as the result of the instantiation of substantive universals by particulars and stating on the other hand that objects or substances represent the instances of substantive universals.

[149] Even if structural *properties* are mentioned at one point: 'a *trope* of the structural property that all methane molecules have is the exemplification of the structure in any particular molecule.' (Ellis 2001: 24, italics added). Compare with 'substantive universals are always instantiated by objects or substances....and properties and relations are always instantiated by their tropes.' (Ellis 2001: 73).

should hold for this (structural) level as well.

In other words, instead of explaining how the instantiation of a substantive universal, *simpliciter*, could justify the presence of 'identical' structures, Ellis is keen to stress only that the other alternatives (on the level of universals) cannot provide the required justification. Ellis seems then to conjecture that, since there needs be something on the level of universals following the instantiation of which kind members should possess the *same* structures, that something must be substantive universals. But surely, rejecting all other alternatives (for having the N&S requirement justified for structures) does not mean that the alternative we are left with should not be explained on its own. Presenting structures as 'substantivised' is helpful in a sense with the intuitive side, but it does not have much argumentative weight.

I conclude that substantive universals cannot justify the N&S and non-overlapping requirements. The substantive universals strategy advocated by Ellis has the same shortcomings as the Lockean strategy, and faces in addition the problem that the nature of the relation between substantive and non-substantive universals is not clarified. Ellis does not solve the difficulty that structural properties pose for his account.

In the next section, I will look at criterion VIII and its possible justifications.

§2.3 The problem of processes and criterion VIII

In the present section, I turn to criterion VIII and the problem of justifying the commonsense phase/non-phase distinction. Criterion VIII asks that natural kind members should not instantiate their kind specific-properties as part of a process in which more fundamental kinds are involved. In other words, it is demanded that those 'divisions in nature' corresponding to natural kinds it be non-phase (i.e. non-implicated

in transitory transformations).[150]

Usually, kind theorists overlook this criterion, as such, but when it is brought into discussion, the reason is nearly always to defend the commonsense phase/non-phase distinction. This intuitive, commonsense distinction leads us to consider certain kinds as depicting divisions of objects, substances and hence being natural, non-phase kinds (and of course, the paradigm here is represented by the kinds in the exact sciences). These are opposed to the kinds that depict stages in processes undergone by objects, which should be considered non-natural, phase kinds (and the paradigm here are certain kinds in biology related to developmental transformations).[151]

Since one common objection against diseases being natural kinds is that diseases represent processes, this criterion deserves separate attention and treatment. Is it justified to rule out medical kinds (and special science kinds in general) as phase kinds? While a more detailed discussion of this aspect will be undertaken in the next chapter, I want to argue here, on a more general level, that without invoking the identity sense of carving nature, criterion VIII cannot be used to rule out special science kinds as phase kinds.[152] I will focus here on what I consider the best philosophical defence in the literature of the commonsensical phase/non-phase distinction, namely Jonathan Lowe's.

What is interesting about Lowe's position is that the diachronic conditions of identity for objects do play a role in defending the common-sense phase/non-phase distinction. These conditions are *not* linked with (non-phase) kind membership, however, but with more general categorizations of objects (in material, artefactual and biological strata). Lowe claims, among others, that unless these conditions of diachronic identity are accepted, the perspective of the number of natural, non-phase kinds multiplying greatly

---

[150] In the notation I have introduced in chapter 1, criterion VIII says that, for any individual *a*, member of a kind K, the determining properties part of the pattern $[p_1, p_2,...,p_n]$ should not represent a *phase* of a more fundamental pattern possessed by *a*.

[151] Almost none of the medical kind (or more generally special science) kind theorists seriously approach this problem.

[152] Certainly, most special science kinds appear intuitively as phase kinds. We should not only think about medical kinds but also economic kinds (discussed in Nelson 1990), biology kinds (discussed in Ereshefsky 1992, 1998), psychological kinds (discussed in Cooper 2008) or even kinds of concepts (discussed in Machery 2005)

above the number of kinds commonly accepted (the so-called 'Heraclitan' threat) and/or reducing greatly below the number of kinds commonly accepted (the so-called 'Spinozan' threat) would be un-escapable. My point will simply be to show that these Heraclitan and Spinozan challenges cannot be repelled by Lowe's strategy.

But what precisely is the connection between the Heraclitan and Spinozan views and the discussion about kinds and their phases? The connection has to do with the transitions, processes or changes undergone by kind members. Simply put, on the Spinozan view almost any transition undergone by kind members amount to a phase kind change, even when it is not intuitively so. On the Heraclitan view, in turn, almost no transition represents such a change – they are almost all non-phase, substantial kind changes. Let me give some examples.

For instance, given the transition undergone by an individual caterpillar turning into an individual butterfly, the Heraclitan view would be that we have a substantial kind change, such that the individual in question changed membership from one natural kind (caterpillar) to another (butterfly).[153] The Spinozan view would be in agreement, in this case, with the common intuitions saying that a phase kind change is in place, such that the individual in question did not change its membership but remained all along part of the butterfly kind. Nonetheless, in other cases, the Spinozan view would also be in conflict with common intuitions. For instance, for the transition undergone by a sample of gold turning into a sample of lead (in a supernova), the Spinozan view would be that a phase kind change is in place as well, such that only one natural kind was involved (either gold or lead) with its *phases*.[154]

On the Spinozan side then, gold would be a phase of lead (just as caterpillars would be

---

[153] The transition from caterpillars to butterflies is part of a set of stock examples employed for expository reasons in the literature on kinds and their phases, when the accent is placed on the metaphysical and not scientific side of the discussion.

[154] I shall adopt in this section a few terminological adjustments, merely for expository reasons. Hence, I shall speak about (natural) 'kinds' and their 'phases', (instead of 'natural kinds' and 'phase kinds'). As 'particulars' will be designated kind members (and not tropes or modes) - where kinds could be countable or mass kinds. Members of countable kinds will be designated as 'individuals' and they will be the main focus of this section. However, everything I shall have to say here about countable kinds is readily applicable to mass kinds as well. It is for this reason that when discussing the absolute identity of kind members, I do not mention their numerical identity but their diachronic conditions of identity.

phases of butterflies). Overall, the number of natural kinds is greatly reduced, because most (commonly accepted)[155] substantial kind changes are not acknowledged, as such. On the Heraclitan side, to the contrary, butterflies and caterpillars would be both natural kinds (just as gold and lead would be two different natural kinds). Overall, the number of kinds is greatly multiplied, because most (commonly accepted) phase kind changes are not acknowledged, as such.

Notably, the abovementioned *kind* changes - phase *kind* changes and substantial *kind* changes - should be differentiated from what we could call phase *individual* changes and substantial *individual* changes. The *individual* changes have to do with the diachronic identity of kind members, *qua* individuals, and not with their *kind* membership. In the case of the caterpillar-butterfly transition for instance, it is one thing to ask whether the individual in question changes its kind membership or not, and another to ask whether the individual remains identical with itself or turn into a different individual. The two questions are related but still different because, for example, it is not impossible to think that the individual caterpillar could undertake a phase kind change while becoming a different individual, that is, while undertaking a substantial individual change.[156] Similarly, no logical entailment exists between substantial kind changes and substantial individual changes – a particular could swap natural kinds while its diachronic identity is not affected.

To be sure, there are variants of the Spinozist or Heraclitan strategy that could be directed at individuals (with the purpose of showing that the number of individuals is dramatically reduced or indefinitely multiplied, respectively). However, in a similar vein, one can affirm that Spinozism/Heraclitanism about kinds does not entail, nor is entailed by Spinozism/Heraclitanism about individuals. For instance, a Spinozist ontology of kinds in which their number is greatly reduced appears consistent with an orthodox ontology of individuals in which the number of kind *members* is not even under dispute. Think of a

[155] Talk of commonly accepted natural kinds in this context is not incongruous with my description of the multi-faceted debate on kind Realism from chapter 1. In this section, by 'commonly accepted kinds' I simply refer to what most lay people prefer to call 'substances', which are sharply differentiated from 'processes' (and the 'phase kinds' that that could be circumscribed by looking at processes). Many scientists, especially the chemists, share the same linguistic intuitions directed at the exact science kinds.

[156] See Geach's famous example of the cat Tibbles *apud* Noonan (2006)

Spinozist, counter-intuitive scenario in which the only accepted kinds were, say, 'inorganic matter' and 'organic matter'. Such a scenario seems consistent with the existence of all the individuals we intuitively accept – their individuation would depend, just as in the more commonsense situation in which we have many other kinds delineated by the natural sciences, on their spatio-temporal locations and on their kind membership. It is just that they would have only two options for kind membership.

As I said, Jonathan Lowe has attempted to defend the commonsense phase/non-phase distinction from the Heraclitan and Spinozan challenges by way of drawing - in an original and elegant argumentative path, which involves substantive universals, laws and categories circumscribing individuals' diachronic identity – a certain dischotomy between *substantial* kind changes and *phase* kind changes.

I will present in the following Lowe's solution (§2.3.1) and then will argue that it is unsatisfactory, mainly because his construal of what kind membership is - the instantiation of substantive universals, *simpliciter* - cannot preclude the "moves' of inflating or reducing the number of kinds in one's ontology. A proponent of the Heraclitan or a Spinozan views could speculate the role that laws play in defining kind changes, while accepting the broad lines of Lowe's metaphysics. I will present the main parts of this law-based strategy (§2.3.2) and then will press on the consequences that this strategy has on the coherence of the ontological picture that Lowe wishes to offer as a foundation for natural sciences (§ 2.3.3). In the final part of this section I will indicate that viewing the diachronic identity of kind members as dependent upon their kind membership could salvage Lowe's position (§2.3.4).

To stress, my purpose is not at all to suggest that it is meaningless to talk about phase kinds. I simply want to say that insofar as criterion VIII is used to rule out special science kinds as phase kinds, then the only chance to do so is by adopting the identity sense of carving nature at its joints. That the alternative of having special science kinds as non-phase kinds entails some form of 'Heraclitanism' does not represent in the least a catastrophic perspective. Such a possibility should simply make us re-think the phase, non-phase distinction, in terms different from the commonsensical ones.

§ 2.3.1 LOWE'S STANCE

Within the construal of kinds as substantive universals, Lowe draws the distinction between substantial kind changes and phase kind changes by way of appealing to natural laws and to his ontological *categories*. Natural laws (which are taken to govern the behaviour of kind members) correspond in Lowe's scheme to relations of 'characterisation' of substantive universals by non-substantive universals.[157] Categories are in turn broad divisions of reality within which the diachronic identity of particulars is confined.

For Lowe, laws and categories stand on the opposite sides of (a certain construal of) the *a priori/a posteriori* dichotomy. Whereas knowing precisely what laws govern the behaviour of kind members depends upon empirical information, categories represent an *a priori* matter, in the sense that they are associated with the metaphysical reasoning that should provide a foundation for the natural sciences (and should reject ontologically unacceptable positions).[158] *Contra*, 'empirically minded' philosophers, Lowe argues that we should *a priori* accept *three* categories (the biological, the material and the artefactual) as the most general divisions of reality, which allow us to track down individuals through change/over time, by being specifically associated, each of them, with certain diachronic conditions of identity.

The rationale is the following: as long as an individual retains its category, all the changes it undergoes represent phase *individual* changes. Any trespassing of its category however (whether physically or only metaphysically possible) could only amount to a substantial *individual* change, associating the coming into existence of a different particular.[159]

---

[157]In Lowe's sortal logic notation, for a particular *a*, a kind ϕ and a non-substantive universal F, 'a/ϕ' represents a relation of instantiation of a substantive universal, laws have the form 'ϕF' and the possession of the property F by the (kind member) *a* should be figured as 'Fa/ϕ'. (See Lowe 1982, Lowe 1989, ch. 8) The non-substantive universal F 'characterises' the kind ϕ in such a way that the identity of ϕ depends on it being characterised by F. (*Cf.* Lowe 2006, ch. 7 and 10, *esp.* pp. 155, 169, 170, 173; see also Heil 2006, p. 7)

[158] In addition, three different senses of the *a priori* (the Aristotelian, Kantian and the British empiricist one) are alluded to by Lowe in various places. In Lowe (1989) for instance, the British empiricist sense (on which *a priori* means *innate*) is touched upon at p. 13 and the Kantian and Aristotelian senses are mentioned as well on pp. 8, 15.

[159]The diachronic conditions of identity should be distinguished (in Lowe) from the *sortal* conditions of identity, the latter being conceptual criteria allowing us to identify and differentiate particulars *as* members of kinds; see Lowe 2001: 59

Indeed, in arguing in favour of the existence of categories, Lowe already tackles the side of the Spinozan and Heraclitan strategies concerning *individuals,* which is also an ontologically unacceptable position.[160]

Now, as I said, Lowe also employs categories and laws in order to face the 'threat' of Spinozism and Heraclitianism about *kinds.* Lowe argues, on the one hand, that substantial *individual* changes also qualify as *kind* substantial changes.[161] On the other hand, for a change to represent a phase *kind* change, it should be the case that the laws that are characteristic of the kind at hand govern the transitions undergone by its kind members:

'A change to an individual substance S, of kind K, is a phase change for S just in case it is a change which things of kind K survive as a consequence of the natural laws of development for Ks' (Lowe, 2001, p. 186)

While in the case of biological organisms we have the natural laws *simpliciter,* in the case of material objects Lowe describes in addition three conditions that the phase changes must meet – they should be reversible, gradual and reproducible.[162] We can infer then - bearing in mind the more general definition of kind changes reproduced above - that for material objects we have a subset of laws in which the members of kinds are involved when undergoing phase changes – the laws governing reversible, gradual and readily replicable transitions.

Lowe's (rather fragmentary) considerations about kind and individual changes, laws, categories, phases and substantial transformations could be systematised as follows. For phase kind changes we have two conditions that are necessary and jointly sufficient, to wit, that the category kind members belong to is preserved and that the laws of the kind

---

[160] Lowe 2001: 86, 187, Lowe 1989: 56-57 *et. passim.*, Lowe 2009: 30

[161] Lowe does not put much emphasis upon the difference between *individual* and *kind* changes. Lowe admits for instance that there are *two* types of substantial changes that particulars can undertake (2001: 174-175). Nevertheless, in certain places in his argumentation it is not very clear which of the two types of substantial changes is discussed, i.e. kind-related or individual-related (see for instance his 2001: 178, 179 *et passim*, 1989: 56-57). Accordingly, it is not very clear which of two possible sides of the Heraclitan and Spinozan strategy – directed at kinds *or* individuals – is criticised by Lowe and attempted to be refuted. Lowe appears at times to conclude that Heraclitanism about *kinds* can be rejected, using premises that could only refute Heraclitanism about *individuals.* It is in order to clarify possible misunderstandings of this sort that I have introduced above a fourfold differentiation between types of changes – with the risk of making my exposition harder to follow

[162] Lowe 2001:176

in question govern the transitions undergone its members. The negation of each condition is in turn sufficient for substantial kind changes. Substantial kind changes occur either when the category kind members initially belong to is not maintained or when the laws of the kind at hand do not govern the respective transformations.[163]

Given the relation between categories and individual changes, we can re-phrase the above conditions by saying that for phase kind changes, phase individual changes and the bearing of laws characterising the kind in question are necessary and jointly sufficient. For substantial kind changes on the other hand, it is sufficient either that an individual substantial change is in place or that the laws guiding the manifestations of the kind in question do not govern the processes its (initial) members come to undertake.

[163] Of course, in Lowe's framework it makes perfect sense to speak about laws characterising kinds and allow at the same time that some of the kind members could undergo transitions that are not governed by those laws. That is because, again, for Lowe laws represent relations between substantive and non-substantive universals, and the law statements are second order quantified. Indeed, the fact that his account offers a metaphysical explanation as to why exceptions to laws' manifestation show up at the level of particulars is taken by Lowe as an advantage of his construal of laws. Otherwise, Lowe argues, we have no way of rendering intelligible the *ceteris paribus* clauses that need to be introduced when analysing the behaviour of particulars in first order quantified statements; see Lowe (1989: 153-154) and Lowe (2006: 13-17)

FIGURE 5 LOWE'S SCHEME

Changes undergone by an individual *a* - initially member of a kind ϕ – which could become member of a different kind ϕ* or go through a phase of ϕ depending of certain necessary and/or sufficient conditions. The conditions are stated in terms of categories and laws (left) and in terms of laws and *individual* changes (right)

Phase kind changes (for ϕ)	The bearing of laws specific for ϕ (necessary)	The preservation of the category *a* initially belongs to (necessary)
	(necessary and sufficient)	
Substantial kind changes (between ϕ and ϕ *)	The laws specific for ϕ are disobeyed (sufficient)	Transgression of the category *a* initially belongs to (sufficient)

The bearing of laws specific for ϕ (necessary)	Phase *individual* changes for *a* (necessary)
(necessary and sufficient)	
The laws specific for ϕ are disobeyed (sufficient)	Substantial *individual* changes for *a* (sufficient)

The way in which Lowe understands kind changes, however, does not fulfil the purpose of responding the Spinozan and Heraclitan challenges of the common-sense phase/non-phase distinction. In the following two-subsection, I will explain why I think that Lowe's view is consistent with an Heraclitan or Spinozan reality of kinds.

§ 2.3.2 THE MAIN STRATEGY

To recall, the challenges that the Spinozan and Heraclitan views pose to the commonsense reality of *kinds* are that many transitions undergone by particulars amount to substantial kind changes, even when they are not intuitively so (on the Heraclitan side) or that many transitions do not represent changes of kind, even when they are intuitively

so (on the Spinozan side).[164] Thus, on the Heraclitan side, the number of kinds in nature is much greater than the one commonly accepted (in exact sciences and/or by laypersons) whereas on the Spinozan side, the number of kinds is much smaller that the one commonly accepted.

Now, is Lowe's scheme up to this task? My answer is - no. To resist, a proponent of the Heraclitian or the Spinozan view (henceforth a Heraclitan or a Spinozan, respectively) would concentrate on the part of Lowe's argumentation in which laws play a role in defining (substantial and phase) changes. Her main strategy would be to question, that is to say, which non-substantive universals characterise which substantive universals, in a sense that will be clear shortly.

Let us look at the changes that are commonly regarded as phase kind changes. These changes are taken to revolve around a *single* kind. Thus, the modification of the features any particular possesses does not amount to the swapping of its natural kind membership for another, but to a kind phase, as in the turning of any particular sample of water into ice or in the transitions undergone by any particular caterpillar along its maturation to a butterfly.[165] The Heraclitan would want to say that no natural laws characteristic of a purportedly *single* kind at work govern such processes and that, to the contrary, we have at least one additional kind being instantiated, a kind which is characterised by its own set of laws. Thus, the Heraclitan would describe the commonly regarded phase kind changes as substantial kind changes.

Let us look now at the changes that are commonly viewed as substantive kind changes. These changes are taken to revolve around (at least) two kinds, such that the modification of the features possessed by a particular amounts to the swapping of its kind membership for another, as in the transformation of any particular sample of gold into

---

[164] Lowe sometimes treats the concerns about individuals and kinds together. In his (2001) for instance, Lowe affirms that one of the advantages of adopting his *kinds* framework is that the Heraclitan view on *individuals* can thus rejected (p. 187) even if, it does appear to me, the *argumentative* role of rejecting it is played in his metaphysics by his three *categories*. See also Lowe's (1989, pp. 56, 57 and 2009, p. 30) where *kinds* (and the sortal conditions of identity associated to kinds) are also invoked in an argumentation that effectively employs *only* categories (and the diachronic conditions of identity) and thus can *only* rebut Heraclitanism about *individuals*.

[165] I shall keep up with the expository examples used by Lowe.

lead for example. The Spinozan would want to say that no natural laws are trespassed and that, to the contrary, the transition is governed by laws characteristic of the initially instantiated kind, such that the putative kind that comes to be instantiated is in fact a phase of the kind initially instantiated.[166] Thus, the commonly regarded substantial kind changes would be described by the Spinozan as substantial kind changes, by way of interpreting (just as the Heraclitan would do, save in an opposite sense), the role that laws play in Lowe's very distinction between substantial and phase changes.

As I said, these strategies would boil down to questioning which substantive universals are *characterised* by which non-substantive universal. That is because, to recall, laws involve for Lowe non-substantive universals characterising substantive universals. Let us take a change in which, say, a particular *a* initially possessing three properties F_1, F_2 and F_3, loses the property F_1 and acquires another property F_4.[167] The question is - what substantive universal(s) would be instantiated by *a*, as a kind member (before, during and) after the change?

Consistently with Lowe's basic account of substantial and phase changes, I claim, one could equally say that:

i) along the transition, *a* instantiated a *single* substantive universal ϕ (characterised by the non-substantive universals F_1, F_2, F_3 and F_4), such that the change was in fact a phase one,

or that:

ii) ii) *two* kinds ϕ and ϕ^* were involved in the respective process, such that the particular *a* changed membership from a kind ϕ characterised by the non-substantive universals F_1, F_2 and F_3 to a kind ϕ^* characterised by the non-substantive universals F_2, F_3 and F_4.

In Lowe's notation, the transition observed on the level of particulars from

[166] or, conversely, the initially instantiated kind could be seen as a phase of the kind that comes to be instantiated.
[167] Where, of course, the properties F_1, F_2, F_3 and F_4 instantiated by *a* would in fact be, in Lowe's metaphysics, modes of the respective non-substantive universals.

$F_1a/\phi_{(?)}\&F_2a/\phi_{(?)}F_3a/\phi_{(?)}$ to $F_2a/\phi_{(?)}\&F_3a/\phi_{(?)}\&F_4a/\phi_{(?)}$,[168] might equally be said to be governed by the law L: $\Phi_{(1)}$ $(F_1\&F_2\&F_3) \rightarrow \Phi_{(1)}$ $(F_1\&F_2\&F_4)$,[169] or the law L*: $\Phi_{(1)}$ $(F_1\&F_2\&F_3) \rightarrow \Phi_{(2)}$ $(F_2\&F_3\&F_4)$.

If the change undergone by *a* was commonly regarded as a phase kind change, the Heraclitan would claim that L* and not L is the law governing it. Hence, a necessary condition for a phase kind changes of the kind ϕ would not be fulfilled and a sufficient condition for substantial kind change to ϕ^* would be satisfied since no laws of the form of L govern the change.

If the change undergone by *a* was commonly regarded as a substantial change, the Spinozan would claim that L is the law governing it. Hence none of the sufficient conditions for substantive kind change from ϕ to ϕ^* would be satisfied and both the necessary conditions for a phase change of the kind ϕ would be fulfilled.

As Lowe himself admits, the definitions of kinds and laws are dependent in his metaphysics.[170] This does not amount to vicious circularity insofar as one is concerned with interpreting law-statements as having kinds as truth-makers and takes laws as describing the relation of characterisation of substantive universals by non-substantive ones. Nevertheless, one can still discuss what and how many kinds (substantive universals) exist out there.

To put it differently, one can accept Lowe's arguments that laws should be viewed as describing characterisation-relations on the level of substantive universals (because otherwise we have no metaphysically unambiguous justification for the nomic necessity, the evaluation of counterfactual conditionals, the fact that laws could have exceptions, etc).[171] Nonetheless, there is still space to discuss precisely *what* kinds are out there, as substantive universals characterised by non-substantive universals.

As we have seen, at least in the first instance, substantive universals can plausibly be

---

[168] or, differently put, the transition observed on the level of particulars from $(F_1\&F_2\&F_3)a/\phi_{(?)}$ to $(F_2\&F_3\&F_4)a/\phi_{(?)}$

[169] or, more generally, the law L: $\Phi_{(1)}$ $(F_1\&F_2\&F_3\&F_4)$

[170] Lowe 1989: 153, Lowe 2006: 16, 127-132, 152-156

[171] Lowe 1989: 150-154, Lowe 2006: 26, 132, 133, 147

added to/retracted from our metaphysical picture so as to modify the number of commonly accepted kinds and challenge our intuitive judgements about phases and substantial changes. Various other arguments Lowe adduces in order to confer intelligibility to his scheme could not preclude, I believe, such a counter-intuitive combining of substantive and non-substantive universals. In other words, I will suggest, the broad lines of his metaphysics could easily accommodate a Heraclitan or Spinozan stance on the reality of kinds.

§ 2.3.3 *Could Heraclitan or Spinozan kinds really be fitted into Lowe's ontology?*

Lowe maintains that judging *normality* (that is, judging which non-substantive universals characterise which substantive universals, irrespective of what properties are instantiated by *particulars*) is not independent of knowing laws[172] (which, of course, describe the relations of characterisation between these two types of universals). But this could not stop the Heraclitan in taking for instance *white ravens* as a non-phase kind, different from the commonly regarded *raven* kind (which has blackness among its central features, according to Lowe), and the change of any particular black raven into a white one (say, as a result of irradiation) as a substantial kind change. What the Heraclitan would claim was that no natural laws concerning the common kind of (black) ravens (which would be commonly taken as the single non-phase kind at work in such a change) are violated, since the black ravens are not supposed to be white. The non-substantial universal whiteness, a Heraclitan would instead contend, characterises the *white* raven kind.

One does not have to stick to examples of laws that have exceptions at the level of

[172] Lowe 1989:153, Lowe 2006: 16

particulars.[173] A Heraclitan would be happy to accept what Lowe says about experiments, for instance. In experiments, after checking that a sample belongs to a certain kind (i.e. that it possesses the tropes of the non-substantive universals that are known up to the moment of the experiment to characterise the kind in question), the behaviour of the sample is tried in hitherto untried conditions. Thus, new laws are derived, often in a *single run* of the experiment.[174] Let us envisage, on behalf of the Heraclitan, an 'experiment' (say, for the people in Jamaica in Locke's times, to use the expository example mentioned by Lowe in his 2001) in which water was subject to below 0°C temperatures. Now, the 'discovered' law that ice/water is solid could be taken by the Heraclitan to characterise the kind *ice*. "*Water* is not supposed to be solid", she would claim – "only *ice* is, and ice represent a different natural kind from water", the Heraclitan reply would go.

On the other side, of course, the Spinozan would not agree with the Heraclitan, since her purpose would be to justify a decreased number of kinds and interpret as phase changes the transitions that are commonly regarded as substantial changes. To the contrary, the Spinozan would hold for instance that an atom (or a macroscopic chemical sample) could possess 79 and then 82 protons as a member of the *same* kind, just as samples of water (as members of the *same* kind) could possess different states of aggregation. Having 79 protons and having 82 protons in nucleus would represent for the Spinozan different non-substantive universals characterising the *same* substantive universal, say gold, and not two substantive universals, gold *and lead*, as it should normally be the case (if Lowe's metaphysics of kinds was to underpin the commonly accepted kinds in the natural sciences). The transformation of any sample of gold into a sample of (what is commonly taken as) lead would be governed in her rationale by the laws characterising the *kind* gold only and hence both the necessary and the sufficient conditions of a substantial kind

[173] I do not have the slightest issue here with Lowe's metaphysical construal of laws according to which they have exceptions on the level of *particulars*. Indeed, it seems to me that in the case of *genuine co-instantiation*, the only way to justify the so-called *ceteris paribus* clauses that are attached to the first-order quantified generalisations frequently appealed to in science is to invoke substantive universals and kinds as truth-makers of laws (see Lowe 1989: 153, 154). I have just taken up the criticism addressed to Lowe that 'normality' is hard to ascertain on the level of particulars (see e.g. Mumford 2000) - a criticism which is I think unjustifiably directed at Lowe's metaphysical account of laws - and have adapted it to it for the Heraclitan's advantage.

[174] Lowe 1989:154

change would not be satisfied.

To be sure, we have the intuition that gold and lead, as elements are distinct and stable, and thus should be regarded as different kinds.[175] Nevertheless, water and ice are also stable on their own, and, as Lowe says, the difference in atomic number is not enough for drawing the distinction between phase and non-phase.[176] Assuming here the role of the Spinozan, I gladly agree.

In the case of material objects, Lowe argues in favour of other intuitions that I have already referred to – the phase changes should be reversible, repeatable and gradual. In other words, the laws specific for a material kind (and what non-substantive universals characterise substantive universals) in its phases, should be picked out using certain other intuitions we have about the transitions they govern on the level of particulars. Nonetheless, the obvious reply would go, intuitions can vary, and a Heraclitan or Spinozan can have intuitions of their own, just as, to be sure, the numerous proponents of an ever changing or everlasting reality in the history of philosophy had. It seems like Lowe would have no answer if in the discussed example, a Spinozan replied that for her it is sufficient that the gold-lead transmutation is realised in both senses in laboratory conditions and occurs naturally and reversibly in supernovas?[177]

One final issue (of consistency) might be worth mentioning here. To return to our example, given the non-substantive universals F_1, F_2, F_3 and F_4 and two substantive universals ϕ and ϕ^*, to ask which F_s characterise which ϕ_s is to question, in a sense, what 'individualises' substantive universals. Lowe does allow exceptions for laws on the level of particulars, with regard to what properties are instantiated by kind members. On the level of universals though, the non-substantive universals should characterise substantive universals as a matter of metaphysical necessity, as it were, in that the identity of substantive universals should depend on the non-substantive universals they are

[175] Even if not all chemical elements are that stable. Unobtainium[346] is an extreme example.
[176] See Lowe 2001:178
[177] Of course, a Spinozan or a Heraclitan could also dismiss the intuitions for regarding, on a commonsensical basis, that certain transitions are 'development changes' governed by 'development laws'. Again, the Spinozan and the Heraclitan could have intuitions of their own.

characterised by, at least in the case of fundamental kinds.[178]

The argumentative pattern involved in Lowe's reasoning for the level of substantive universals (that is, in the case of *kinds*) is striking similar to the argumentative pattern employed by anti-Humean proponents of the metaphysical necessity of natural laws (in the case of *properties*).[179] The latter theorists claim that (some or all of) the non-substantive universals are individualised by the causal powers they associate, in that the identity of non-substantive universals depends on the causal interactions their instances can engage in.

Lowe criticises this approach, mainly on the ground that it involves a confusion between principles of *trans*world identity and principles of *intra*world identity.[180] That is, it involves drawing (unwarranted) conclusions from what is *actually* the case to what is metaphysically necessary the case, along the path of Kripke's *a posteriori* necessities.[181] Lowe claims therefore that the respective approach has two problems. It has first a strong, metaphysical problem – to suppose that the identity of a non-substantive universal depends over its association with powers just seems 'highly dubious.'[182] It also has a weaker, epistemological difficulty in that one cannot be sure about which non-substantive universals are individuated by which powers.[183]

Now, referring strictly to the weaker, epistemic issue, my suspicion is that Lowe has a problem of consistency here, because his main stance on how kinds are identified and 'singled out' is simply that in nature we find certain *stable* co-instantiations of properties.[184] This line however appears as highly susceptible to the same type of critique Lowe advances against the anti-Humean proponents of nomic metaphysical necessity. It is *actually* the case that on the level of particulars, properties are agglutinated in certain

[178] See Lowe 2006: 155, 169, 170, 173. Lowe avoids the of terminology of identity-dependence for substantive universals, preferring to say instead that (at least) some laws concerning natural kinds, when properly *interpreted*, emerge as metaphysically necessary (*ibid.* p. 155).
[179] See for instance the arguments advanced by Shoemaker 1998, Ellis 2001, Bird 2005
[180] *Cf.* Lowe 2006: 150-152,164-165. Several other reasons are adduced against the metaphysical necessity of laws, *as* construed by the anti-Humean theorists. For instance, Lowe claims that universal physical constants could have been different (*ibid.* p. 151).
[181] *Ibid.* p. 153
[182] *Ibid.* p. 164-165
[183] *Ibid.* p. 154
[184] *Ibid.* p. 135.

patterns that science informs us about, one could reply. Why is it warranted to draw conclusions, on such empirical grounds, about which non-substantive universals characterise which substantive universals, where the relation of 'characterisation' involves the latter's identity?[185]

I do not wish to question here Lowe's view of the nature of the 'characterisation' relation. There might be independent reasons why the respective relation needs to be construed in this way, especially in order to differentiate it from Armstrong's relation of necessitation.[186] What I strongly suggest though is that when Heraclitanism or Spinozism are reckoned with, Lowe's considerations on laws, kind 'individualisation' and empirical research cannot count as means to resist these views. If to construe the special science kinds as non-phase kinds is a form of Heraclitanism, Lowe's defence of the commonsense phase/non-phase distinction cannot count as a rejection.

As in the discussion of criteria VIb and VII, I want to bring into the picture the identity sense of carving nature at its joints. I will show in the next section that only if the diachronic conditions of identity for kind members are taken to depend upon their kind membership, the Spinozan and Heraclitan challenges could be responded and, in particular, criterion VIII could be used in order to rule out special science kinds as phase kinds.

§2.3.4 *IDENTITY AND THE PHASE NON-PHASE DISTINCTION*

If the identity sense of carving nature at its joints was adopted, the Heraclitan and Spinozan challenges could be adequately responded. Within Lowe's system, adopting this identity sense as a solution would mean treating the commonly accepted natural kinds as

[185] See especially Lowe 2006: 170, where Lowe rejects the view that the identity of a non—substantive universal might depend on its association with a substantive universal (using the example of electric charge). On the same page, Lowe intimates that kinds are 'individualised' by certain non-substantive universals they are characterised by (the example of electrons, positrons and unit charge). In plain words, for Lowe, to speak of the identity of non-substantive universals (either vis-à-vis their association to kinds, or their association to causal powers) is ill-founded, whereas talk of the identity of kinds is justified. No reason for this unilateral treatment is provided.

[186] Armstrong (1997)

categories. That is because, given Lowe's construal of what a category is, this move would *inter alia* entail that the identity of kind members depended on their kind membership (such that the diachronic conditions of identity for any individual would have to include and mention explicitly the instantiation of a substantive universal).

Recall that in one formulation of Lowe's basic scheme on changes (which I have framed in § 2.3.1). Individual phase changes and the bearing of laws characterising a kind should be necessary and jointly sufficient for the changes undergone by its members being phase ones. On the other hand, for substantial kind changes it should be sufficient either that an individual substantial change is in place or that the laws governing the manifestations of the kind in question do not have a bearing over the respective processes undergone by its members.

Now, if the commonly accepted natural kinds were treated as *categories*, not only all individual substantial changes would qualify as kind substantial changes (as already is the case in Lowe's scheme) but also all substantial kind changes would be individual substantial changes. That is, individual substantial changes would be sufficient *and necessary* for kind substantial changes. This would eliminate from Lowe's scheme the role that laws play in defining substantial kind changes and would stop the Heraclitan and/or the Spinozan from seeing such changes where intuitively they are not and/or overlooking them where they intuitively they are in place.

The role played by laws in *phase* kind changes would have to be maintained though, for the following reason. There are some changes undergone by particulars that, intuitively, are neither phase changes nor substantial kind changes – like for instance in the case of the irradiation of a raven that turns thus white. Lowe's construal of laws coherently allows such exceptions on the level of particulars. This is in fact one of the great advantages of his account of natural laws. Notice, however, that on his scheme of changes - which includes laws playing a role in the definition of *both* substantial and phase kind changes - the process in question could not qualify as a phase change (since a raven turning white would not obey the law of 'development' for the kind of raven). On the other hand, this process could qualify as a substantial change on the other hand. As we have seen, the

Heraclitan would gladly speculate this latter aspect.[187]

Among others, the purpose of viewing the commonly accepted kinds as categories would be to stop the Heraclitan from re-interpreting what substantial kind changes mean. The purpose would also be to defend commonsense intuitions on the phase/non-phase distinction. If the role of laws were rejected altogether from drawing this phase/non-phase distinction, we would have to accept that since the individual ravens plausibly remain identical with themselves if they turn white, changes such as the ones they suffer under irradiation represent phases undergone by their *kind*. Yet, while processes such as irradiation saliently qualify as individual phase changes, our common intuitions say that we should not have *all* the individual phase changes entailing that kind phases are in place. To put it simply, what the commonsense regards as accidents should not count as kind phases.

[187] And, of course, the Heraclitan who wanted to view medical kinds as non-phase, natural kinds would *very* gladly consider this possibility.

FIGURE 6 COMMONLY ACCEPTED KINDS AS CATEGORIES

An individual *a* - initially member of a kind ϕ - could become member of a different kind ϕ*, could go through a phase of ϕ or could come to instantiate an accident vis-à-vis ϕ, depending on certain necessary and/or sufficient conditions. The conditions are stated in terms of categories and laws (left) and in terms of laws and *individual* changes (right)

Phase kind changes (for ϕ)	The bearing of laws specific for ϕ (necessary)	The preservation of the category *a* initially belongs to (necessary)
	(necessary and sufficient)	
Substantial kind changes (between ϕ and ϕ *)	Transgression of the category *a* initially belongs to (sufficient)	
Accidents for ϕ (neither phase nor substantial changes)	Disobeying of the laws specific for ϕ (necessary)	The preservation of the category *a* initially belongs to (necessary)

The bearing of laws specific for ϕ (necessary)	Phase individual changes for *a* (necessary)
(necessary and sufficient)	
Substantial individual changes for *a* (sufficient)	
Disobeying of the laws specific for ϕ (necessary)	Phase individual changes for *a* (necessary)

As I said, the main upshot of adopting the identity sense of carving nature at its joints would be to stop the Spinozan and the Heraclitan from construing in a particular way the conditions for substantial kind changes. *Contra* the Spinozan, one could argue that any transition undergone by an individual changing its characteristics so as to swap two commonly regarded kinds fulfils the necessary and sufficient condition for substantial kind change (since it also represents an individual substantial change). Similarly, *contra* the Heraclitan, one could hold that any transition that does not involve the swapping of two commonly regarded kinds does not fulfil the necessary and sufficient condition for substantial kind change and hence it simply represents a kind phase or a simple individual phase (an accident).

Strictly referring to the Heraclitan challenge, the multiplication of kinds would also be limited by the fact that an individual could not simultaneously instantiate two natural

kinds (that were not in a species-genus hierarchy), since this would violate the *tertium non datur* law. This prohibition is plainly justified when it comes to categories – no particular could simultaneously instantiate two different categories, and Lowe argues very convincingly in favour of it for the statues/materials cases, in which we have different particulars with the same spatio-temporal location.[188] The prohibition in question could easily be extended for natural kinds, if categories were more broadly viewed so as to enclose natural kinds.[189]

Should the two Heraclitan and Spinozan still want to modify the number of kinds by excessively inflating or reducing it, they would have to look at the conditions of individuation for kind *members* and claim that on the level of *particulars* almost every change is a change in their diachronic (absolute) identity (on the Heraclitan side) or that almost no change is a change in the diachronic identity of individuals (on the Spinozan side). However, this would be a much more difficult enterprise, or, at any rate, it would represent a different discussion.

There are some serious difficulties with taking such a view on the identity bearing of kind membership. Some of the most important are the general epistemological problems of the identity sense of carving nature, which I will discuss in the next chapter. Parenthetically, I should add that that the view of the commonly accepted natural kinds as categories would *not* suffer from the tribulations that Lowe imputes to it, in some (again, rather fragmentary) remarks he makes on the putative difference between kind substantial changes and individual substantial changes.

Firstly, this construal of categories and the identity of kind members would not *entail* that every commonly regarded phase individual change was in fact a substantial individual change.[190] I do not see why it should even favour this view (although it certainly does not preclude it). Taking the commonly regarded natural kinds in science as having a bearing on identity is something that should counter the Heraclitan or the Spinozan who pursued the easier (i.e. less implausible) side of their strategy, that is, who

[188] See Lowe (1995) and also Oderberg (1996)
[189] Which is just to say, of course, that on the identity sense of carving nature, criterion VII could also be used to stop the multiplication of kinds
[190] *Contra* Lowe 2001: 187

questioned the commonly accepted *kinds*, and not *particulars*. Should anyone want to argue in favour of a Heraclitan or Spinozan reality on the level of particulars, that kinds have an identity bearing would not matter, I think (or it would constitute a hindrance), since the argumentation would simply have to be pursued on a different level.

At best, one could advance the following inference indicating the dangers of seeing the instantiation of substantive universals as part of diachronic conditions of identity: *if* kinds have an identity bearing, and *if* kinds are indefinitely multiplied (or reduced) *then* an ever-changing (or un-changing) reality of particulars will result.[191] Nevertheless, all hinges on the accent with which this conditional is employed *reductio ad absurdum*, and for what purpose. One could use this conditional to argue that:

i) Kinds should not have an identity bearing because such and such consequences over the reality of particulars emerge,

or

ii) The number of kinds should not be multiplied or constricted because such and such unacceptable consequences about the reality of particulars emerge.

To be sure, Spinozism or Heraclitianism about particulars is much more unacceptable than Spinozism or Heraclitianism about kinds. After all, many respectable metaphysical schemes in contemporary metaphysics do not have a central place for kinds (or do not even acknowledge their existence) but pay full heed to the existence of particular

[191] In fact, I am inclined to reject even this inference, simply because, as I have mentioned in the introduction to this section, it seems to me that Heraclitanism/Spinozism about kinds and Heraclitanism/Spinozism about particulars represent *different* aspects of these counter-intuitive strategies, and these aspects do not presuppose each other. Take for instance one strong Spinozan scenario on *kinds* I have mentioned in the beginning of, according to which our ontology contains only one or, let us say, three kinds, corresponding to Lowe's three categories – the biological, the artefactual and the material. This Spinozan view would be consistent I believe with a construal of kinds as substantive universals and would not have *a fortiori* consequences over the differentiations between particulars we commonly accept. All the non-substantive universals that in Lowe's current scheme characterise various substantive universals in the biological, inorganic and artefactual realms would characterise (what we could call) the *biological* substantive universal, *material* substantive universal and *artefactual* substantive universal, respectively. All the particulars that in Lowe's current scheme are members of various natural kinds would be members of these three kinds only. The differentiations between these particulars would be grounded, as they are grounded already in Lowe's current approach, on certain category based diachronic conditions of identity (which, in this case, would coincide with the sortal conditions of identity).

substances.[192] All in all, there is no relation of *entailment* between viewing the kind membership as having as bearing over the diachronic identity of kind members and all phase individual changes ending up being also substantial individual changes.

Second, on this construal, categories would not be the most general forms of being we could think of.[193] However, that the material, biological and artefactual are the most general categories we could envisage answers only one half of the question concerning their number – why should not we have less than three categories. It does not answer why we should not have more.

Third, this construal would keep open the possibility of kind swapping (like in the case of any organism that dies) but it would not preserve the possibility of an individual changing its kind while remaining the same.[194] This is just to say that it would not admit its logical negation. Why would one want to preserve the possibility of such changes between kinds of numerically identical individuals? The only reason I can think of is the doctrine of Christian transubstantiation. But there are alternative ways to deal with this issue that theologians have formulated – approaches with sidestep the medieval Aristotelianism and interpret in a different sense how Christ is 'present' or 'in place' in the bread and wine of the Eucharist.[195]

Fourth, on this construal we would not be 'born' with the criteria of diachronic identity of objects and we would sometimes rely on scientists to tell us what kinds are out there.[196] Would this mean that categories are not any more *a priori*? I suppose that they would still be, if in a minimal sense, to attribute to kinds identity bearing over their members would not be something that scientists could tell us about, or any 'empirical' information could provide us, but would be the result of a metaphysical reasoning designed to reject certain

---

[192] See for instance Armstrong (1997), Campbell (1990). As we know, even for Quine the preference to 'desert landscapes' stopped at the level of the first order quantification (destined to reduce all the other statements about 'intensional' entities); see Quine (1968)

[193] As Lowe claims they should be; see Lowe (1989, ch. 7). The same claim can be found in Wiggins (2001) with regard to the 'ultimate sortals', which correspond in Wiggins's system to Lowe's 'categories'.

[194] Whereas Lowe claims that this possibility should be preserved (without specifying any reason for this); *Cf.* Lowe (2001: 55, 184), (1989: 14)

[195] Hemming (2000) is a brilliant example.

[196] *Pace* Lowe 2001: 187, for instance, where it is claimed that categories should be somewhat innate (in connection with the British empiricist construal of the *a priori*).

unacceptable positions.[197] The role of metaphysics, as underlying and providing a framework for the natural sciences – a role that Lowe insists upon in various places[198] – would be unaltered.

Of course, these are details that concern the peculiar aspects of Lowe's system rather our general purpose in this chapter. However, Lowe's arguments are important for the discussion of criterion VIII because they represent, as I have mentioned, the best defence in the literature of the commonsense phase-non phase distinction. The general argumentation I have adopted in this section against criterion VIII ruling out special science kinds as phase kinds will be complemented in the next chapter by a more contextualised analysis of medical kinds. In fact, all the arguments presented in the present chapter will be used as a general metaphysical background in chapter 3, where the comparison between gold and the Graves' Disease kind will be undertaken. At the end of chapter 3 the epistemological difficulties of the identity sense of carving nature will also be presented and the metaphysical considerations from the present chapter on essences, substantive universals and processes will be rounded up with my conclusions about medical kinds in particular and natural kinds in general.

[197] Nothing in the reasoning concerning categories entails that have them from birth or even that we should be able to know them in any direct way. Taking categories as 'innate' comes dangerously close to evolutionary biology/psychology views, with all their tribulations, even if Lowe, in his philosophy of mind, (distances himself from and) criticises the pre-eminence that these evolutionary approaches have come to have; see Lowe (2000). More generally, categories are highly susceptible to 'arbitrariness' and 'subjectivism' accusations, if we construe them either as innate or (even) as Kantian transcendentals. The minimal sense of *a priori* associated with the metaphysical reasoning that should be employed in order to reject unacceptable positions (and provide a foundation for natural sciences, as Lowe wants its metaphysics to be) is consistent with scientific information being used in non-empirical metaphysical considerations.

[198] See Lowe 2006: 3-5 *et passim*.

CHAPTER 3. GOLD VS. THE GRAVES' DISEASE KIND

INTRODUCTION

Are diseases natural kinds? I have noted in chapter 1 that, since kinds in the exact sciences (are taken to) typify natural kinds, a straightforward way to address such a question is to compare a kind of disease with an exact science kind. Such a comparison should make clear certain resemblances and differences between medicine and the exact sciences that should matter decisively in establishing the character of kinds of diseases. In the present chapter, I will compare Graves' Disease, taken as an example of a 'pathological' kind and gold, as an example of a 'classical' natural kind. Before explaining from what point of view this task will be undertaken, let me first draw the background image of the comparison.

Imagine we have, on the one hand, a set of various things—say rings, coins, frames, etc.—that are yellow, fusible, malleable, soluble in *aqua regia* and so on. On the other hand, think of a series of organisms that present certain biological features in common—exophthalmia (the protrusion of one or both eyes), myxoedema (a tibial non-pitting oedema), goitre (an enlargement of the thyroid gland detectable as a swelling in the neck), and decreased weight, accelerated intestinal transit, tachycardia, intolerance to heat, osteopenia, osteoporosis, proximal muscle weakness, etc.

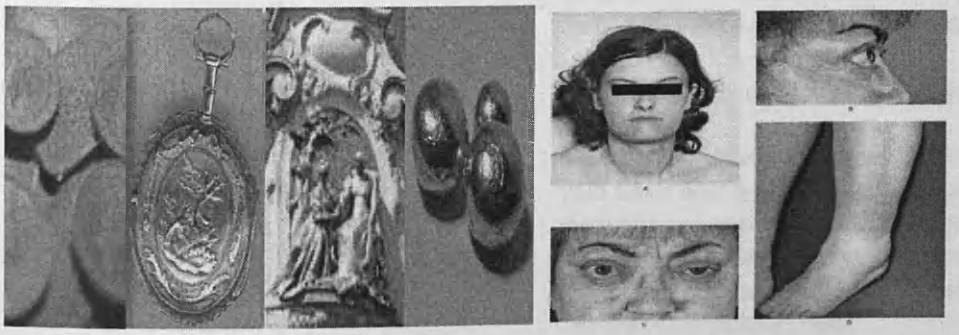


FIGURE 7 GOLD KIND MEMBERS AND ORGANISMS SUFFERING FROM GRAVES' DISEASE (REPRODUCED FROM WHEETMAN, 2000)

Each of these two sets of particulars has certain striking resemblances. The question I am concerned with can then be usefully re-formulated as follows: if gold objects are part of a *natural* kind, why should not organisms with the symptoms of Graves' disease form such a kind as well?

Do the two kinds fulfil/fail the criteria of membership laid out in chapter 1 (and reproduced below), in such a way that an ontological gap between them is salient?

***FIGURE* 8 The criteria of membership in a natural kind NK for a particular *a* instantiating a pattern [p1, p2,.....pn] of 'profound' properties and a pattern [q1, q2,.....qn] of 'superficial' properties**

I. Instantiation of natural properties (natural properties criterion)	I. $[P_1, P_2, \ldots P_n]$ should be natural properties
II.* Instantiation of explanatory powerful, determining properties (causal induction criterion)	II. $[P_1, P_2, \ldots P_n]$ should be responsible for $[Q_1, Q_2, \ldots Q_n]$ (as well as for the causal interactions with other individuals). Were *a* to be NK, a would, or probably would, instantiate $[Q_1, Q_2, \ldots Q_n]$
III. Instantiation of nomic properties (law involvement criterion)	III. $[P_1, P_2, \ldots P_n]$ should be involved in laws (either governing/describing the causal interactions whose effects are $[Q_1, Q_2, \ldots Q_n]$ or the interactions with other individuals
IV.* Instantiation of identity-bearing properties for the kind members (identity criterion)	IV. $[P_1, P_2, \ldots P_n]$ should be involved in the (absolute) identity of *a*. Were *a* to stop instantiating any of the $[P_1, P_2, \ldots P_n]$ properties, it would not be the same numerically identical individual (or it would not have the same diachronic identity conditions, if *a* was a sample of a non-countable kind)
V. Co-instantiation induction requirement	V. The co-instantiation induction criterion – were *a* to be NK, it would instantiate $[P_1, P_2, \ldots P_n]$
VIa. Cluster of properties requirement	VIa. $[P_1, P_2, \ldots P_n]$ form a cluster due to the causal structure of the world; neither all of $[P_1, P_2, \ldots P_n]$ nor any part of them are both necessary and sufficient for membership
VIb.* N&S requirement	VIb. $[P_1, P_2, \ldots P_n]$ are necessary and sufficient for membership
VII. Non-overlapping requirement	VII Was *a* to be a member of a different natural kind NK', then either NK' would be identical with NK or the two kinds would be in a species-genus hierarchy
VIII. Non-phase requirement	VIII. The instantiation of the pattern $[P_1, P_2, \ldots P_n]$ should not represent a phase in a process whose end result is another, non-phase pattern.

* The marked criteria are associated with a construal of properties as *essential*, in different senses though – explanatory essential (criterion II), identity bearing for kind members (criterion IV), identity bearing for kinds (criterion VIb)

Take criterion I, concerning natural properties. Both gold and the Graves' Disease kind (henceforth GD kind) fulfil this requirement, or at least, no ontological gap can be detected at its level. With respect to gold, the resemblances between its members do appear as genuine - they are malleable, soluble in *aqua regia*, yellow, can be extended and shaped without cracking, dissolvable in a HCl/HNO_3 solution, able to reflect the 570-580 nm spectrum of the visible light, etc. These are not properties resulting from wiggling predicates but are pointing towards real similarities, behaviours, and dispositions. being malleable, soluble in *aqua regia*, yellow, being extended and shaped without cracking, dissolving in a HCl/HNO_3 solution, reflecting the 570-580 nm spectrum from the visible light, etc, are not properties resulting from wiggling predicates but point towards real similarities and behaviours. With respect to the GD kind, the observable similarities shared by its members do appear as natural properties as well. Values are not involved in ascertaining whether an organism presents enlargement of the thyroid gland, protruded eyes, a non-pitting oedema of the tibial tissues etc. In the same vein, no gerrymandering of predicates is present in attributing to human organisms properties like exophtalmia, goitre, myxedema, etc. That is, such organisms are not the focus of our discussion because a predicate like, say 'being hospitalized between 7 and 19 days in endocrinology clinics and being either male or female' has them in its extension. These organisms do present biological resemblances (even if they are, from a medical point of view, superficial, in the sense in which medical research is trying to discover the 'profound' causes of such symptoms and signs).

Admittedly, one might invoke here the problem of the theory-ladeness of observation, or the fact that the properties in question might be relational. But ever since the early logical empiricist program of providing a purely observational vocabulary fell into disrepute, the fact that our theories might inform our observational terms is no longer considered to be a serious problem.[199] At any rate, if it is a problem at all, it affects both kinds under discussion. Furthermore, if the charge of sharing relational properties is raised against the GD kind, it can be equally raised against gold. Thus, either both kinds fulfil this

[199] Cooper (2005: 81-86) has argued persuasively that these worries should be dismissed in the case of psychiatry and I take it that the same goes for the case of somatic medicine.

requirement or no ontological gap can be justified at its level. [200]

In §3.1-§3.5 I will take each of the remaining criteria of membership in turn and will argue that instead of an ontological gap between the two kinds, we have a difference of degree. This applied enquiry will follow the general metaphysical discussion from chapter 2 from two (inter-connected) points of view. On the one hand, it will rely on the previous analyses of essences, substantive universals and processes. On the other hand, it will assume – with regard to the intuitive assumption(s) that natural kinds provide a 'basic ontology', frame 'ultimate classifications', tell us 'what things are' etc. - that gold carves nature at its joints in a sense that does *not* involve the identity of its members. The other, identity-related sense will be discussed in the final part of this chapter, in §3.6. Evidently, certain references to the identity of kind members will be present throughout the entire chapter, in connection with my working assumption that theorists appeal tacitly to the identity sense of carving nature when setting forth certain requirements (and characterisations of the natural kind category) – requirements that are otherwise groundless.

§3.1 INDUCTION

Induction has two variants – a causal and a co-instantiation one. In the framework I have adopted, the former has to do with inferences about superficial properties while the latter is expressed by inferences about determining properties.[201] If successful inductions can be framed for the members of a certain kind, the kind in question is a more probable candidate for being natural. Most neutrally put, inductively rich kinds go one step further away from being simply arbitrary collections of particulars and are more probable candidates for circumscribing 'divisions of nature' that correspond to natural kinds. I have called the requirements that natural kinds should participate in these two types of

[200] The same type of argument can be advanced for the microstructural similarities shared by the members of each kind.

[201] The above distinction between causal and co-instantiation inferences does not exclude that, ultimately, causation might also be the metaphysical ground for co-instantiation. Indeed, I shall argue myself in favour of this possibility. These two types of inferences need to be distinguished however, because there are alternative metaphysical justifications for co-instantiation on offer (essences and substantive universals); see also §1.1.

induction criterion II and V, respectively. I shall argue in this section that both gold and the GD kind satisfy them.

3.1.1 THE CAUSAL INDUCTION REQUIREMENT (CRITERION II)

Let us first look at criterion II – the causal induction requirement. Criterion II says that a group of particulars - identified *prima facie* due to their instantiation of certain 'superficial' properties and behaviours [q_1, q_2, ...q_n] - form a non-arbitrary kind K if they also possess certain properties [p_1, p_2, ...p_n], which are 'profound' from a causal point of view, i.e. are attended by causal powers, have roles in scientific explanation, prediction, etc.[202] If this is the case, then for any individual *a*, one could advance inferences about its 'superficial' properties of the form 'Were *a* to be a member of K, *a* would, or probably would, instantiate [q_1, q_2, ...q_n]'.

Gold fulfils this requirement with a vengeance. One of the reasons why grouping together the objects that are malleable, soluble in *aqua regia*, yellow, etc., is not conventional, is that one can reasonably expect that the next observed sample of gold will be extended and shaped without cracking, will dissolve in a HCl/HNO_3 solution, will reflect the 570-580 nm spectrum from the visible light, etc. Such causal inductive inferences can justifiably be framed because, besides these 'superficial' properties and behaviours, gold members also share determining properties – namely the micro-structural properties characteristic of the gold element (and its metallic complex); the latter properties are causes for the former. That is, the superficial similarities of objects with such and such colour, solubility, etc. are indicative of natural kind relationship vis-à-vis criterion II because they are doubled by 'profound' similarities at the level of the micro structural

[202] Thus, from the point of view of the natural kinds discussion, the causal induction requirement differs in extension from the natural property requirement. Of course, one could argue, in line with the Eleatic principle, that a property exists only insofar as it is attended by causal powers. However, the causal induction requirement is particularly concerned with the sort of properties science tries to discover - the determining properties. Even if, for instance, the superficial, perceptible properties of water may also be attended by causal powers (and hence can act as causes) science is interested in the underlying, microstructural properties of water (which act as causes for its superficial, perceptible properties). This is not to say that determining properties need only be micro-structural; it is just to say that, even if all properties may act in some respect as causes (and hence be attended by causal powers) some of these causes are more important than others.

properties of the gold element.

There are, to be sure, numerous questions to be asked about the nature of determining properties (i.e. the individuation of causes, the ontological dependency of causal powers, etc).[203] However, these difficulties about the exact characterisation of the causative relation between the determining and the superficial aspects of kind members nowithstanding, the similarities present on the level of micro-structures do mark out an important distance between utterly arbitrary collections of individuals with different underlying properties and the kinds for whom causal inductions concerning superficial properties can be formulated.

Now, one can affirm that the same is the case for the GD kind. The set of symptoms of Graves' Disease are not grouped together fortuitously, from the point of view of criterion II. That is, one of the main reasons for why 'classificatory strangeness' is not present when considering together the organisms that possess the above-mentioned 'superficial' properties is precisely the common causal ground underlying them.

Indeed, we have here a striking example of a conjunction of symptoms whose causal base has been discovered by modern medicine and thus lies in sharp contrast with disease and disease kinds in the classification of which causal arbitrariness was the rule. The history of medicine is rife with such examples - conditions in which various symptoms were grouped together for reasons that did not have anything to do with the similarities present on the level of determining properties. In the sixteenth century for instance, blindness, chancre and penile 'rubbery tumours' (granuloma formations) were considered syphilis symptoms and their emergence was attributed to astrological influences punishing fornication. In antiquity, to take another example, convulsions, loss of consciousness and vertebral muscle spasms were grouped together as common symptoms of a 'divine' disease affecting certain chosen people, a disease which should have allowed its sufferers to transcend the sublunar realm and enter the realm of essences.

As I said, GD is a different sort of disease, however, because determining 'pathological'

---

[203]See Schaffer (2007) and Psillos (2002) for discussion. What I just want to stress is that my use here of determining properties (as properties attended by causal powers) is neutral vis-à-vis the Humean/anti-Humean dispute, which will be looked at more closely in chapter 4.

properties are also shared by kind members. The diagnosis of Graves' Disease is established by detecting in the blood certain types of antibodies called TRA_s (thyroid receptor antibodies) and high levels of thyroxin (the thyroid hormone). These in turn represent parameters indicating the presence of specific biological micro-structural properties - e.g., increased permeability of cells to sodium and sugar, the presence of certain steroid transcription factors in nucleus, increased Na^+/K^+ ATP_{ase} concentration in membrane, an increase in the number and size of mitochondria, reduced expression of the thyrotropin β subunit and common α subunit (suppressed thyrotropin), increased expression of HCN2 (hyperpolarization-activated cyclic nucleotide-gated catiion channel-2), voltage gated potassium channel and SERCA (sarco(endo)plasmic reticulum calcium-activated ATP_{ase}, responsible for increased heart rate and contractility), increased type 1 5' deiodinase, increased LDL and VDRL receptor lipase (responsible for reduced total and LDL cholesterol), increased osteocalcin, alkaline phosphatase, and urinary N-telopeptide (responsible for osteopenia, osteoporosis, and fractures), increased serum creatine kinase (responsible for proximal muscle weakness and easy fatigability), increased fatty acid oxidation and sodium potassium ATPase (responsible for increased thermogenesis and oxygen consumption, perspiration and weight loss), etc.

These are determining properties because they are each in turn involved as causes in different stages of the global process which underlies the disease in question and whose 'outcome' is precisely the emergence of the symptoms. Some of these determining properties are summarised below, in parallel to their effects.

FIGURE 9 DETERMINING PROPERTIES (reproduced from Brent 2008)

Table 1. Manifestations of Graves' Disease.*

System	Clinical Finding or Manifestation	Marker of Direct or Indirect Thyroid Hormone Action
Pituitary	Suppressed thyrotropin	Reduced expression of thyrotropin β subunit and common α subunit
Cardiac	Increased heart rate and contractility	Increased expression of HCN2, voltage-gated potassium channel (Kv1.5, Kv4.2, Kv4.3), and SERCA; increased α-MHC and decreased β-MHC expression; increased serum atrial natriuretic peptide
Hepatic	Increased peripheral T_3 production; reduced total and LDL cholesterol, lipoprotein(a)	Increased type 1 5′-deiodinase, LDL and VLDL receptor, lipase, SREBP-2, CYP7A, and CETP
Skeletal	Increased bone turnover, osteopenia, osteoporosis, and fractures	Increased osteocalcin, alkaline phosphatase, and urinary N-telopeptide
Reproductive		
Male	Erectile dysfunction, reduced libido	Increased sex hormone globulin, reduced free testosterone
Female	Irregular menses	Antagonism of estrogen action; impaired gonadotropin regulation
Metabolic	Increased thermogenesis and oxygen consumption	Increased fatty acid oxidation and sodium–potassium ATPase
White fat	Reduced fat mass	Augmented adrenergic-mediated lipolysis
Muscle	Proximal-muscle weakness, easy fatigability	Increased SERCA activity and serum creatine kinase
Thyroid	Increased thyroid secretion of T_3 and T_4	Increased type 1 and type 2 5′-deiodinase activity in thyroid

* Data are from Motomura and Brent,[9] Brenta et al.[10] and Klein and Ojamaa.[11] CETP denotes cholesterol ester transfer protein, CYP7A cholesterol 7 α-hydroxylase, HCN2 hyperpolarization-activated cyclic nucleotide-gated cation channel 2, LDL low-density lipoprotein, MHC myosin heavy chain, SERCA sarcoplasmic reticulum calcium-activated ATPase, SREBP-2 sterol regulatory element–binding protein 2, T_3 triiodothyronine, T_4 thyroxine, and VLDL very-low-density lipoprotein.

Because the GD kind presents such an underlying layer of determining properties, it can be involved in (causal) inductions in the same relevant sense as gold. That is to say, criterion II cannot circumscribe an ontological gap at this point. We have instead a difference of degree.

The difference in degree, in this particular context, refers to the certainty of the respective inductive inferences. One (gold) is arguably universal, whereas the other admits exceptions (or, alternatively, can be read as probabilistic) when it comes to the organisms

that are members of a GD kind.[204] In the case of gold, we can expect that the next observed member of this kind *will* dissolve in *aqua regia* for instance, whereas in the case of GD we can expect that the next observed member of this kind is to have exophthalmia with a certain probability. Alternatively, we can expect that the respective organism *will* have exophthalmia only if a host of certain other conditions are in place.

However, such a difference in the certainty of inferences does not amount to the two kinds being situated on different ontological levels. Insofar as the requirement to participate in causal inductions is concerned, a kind is inductively rich if causal powers are at work between the involved properties (i.e. the superficial and determining ones).[205]

Irrespective of the disagreements as to what interpretation causal powers should receive and the exact nature of determining properties, at this point we only need to stick to the (minimal) fact that causation exists in medicine. In other words, if in pathological organisms there are non-accidental successions between events, inductions should be vindicated and criterion II is fulfilled (even if the respective inferences are not universal).

Take the relation between TRAs and exophthalmia, a relation that is considered by all of today's medical standards as causal. This relation has a known mechanism;[206] it has been verified in experimental conditions, both *in vivo* and *in vitro*, and has been subject to

---

[204]More precisely, the respective inferences can be read as probabilistic (on the frequency-interpretation of probability), in which case the inductions simply express the chances of certain properties designated in the consequent to show up given the presence of other properties designated in the antecedent (and thus the respective inferences are metaphysically underlined by non-deterministic causation); see Forster (2004: 21,54). The exceptions, on the other hand, can be taken as showing up due to the presence/absence of certain other factors that are not specified in the antecedents of inductive conditionals (and in this case, the 'probably' locution present in the inductive statements would have an epistemic interpretation); see e.g. Fales (1990: 105). Alternatively, the other factors that might influence the conditions in which the antecedent of the respective conditionals obtains might be indicated in an attached (ceteris paribus-like) clause; see for instance Cartwright (1999:90). In medical treatises (endocrinological ones included, of course) the correlations between causative factors and effects are sometimes expressed as frequencies *simpliciter*, but the *addendum* that certain other factors, currently unknown, might influence the strength of association is just as often met; see for instance Bahn *et. al.* (1998)

[205] See Morris's (2007) presentation of Hume's inductive scepticism related to causal *powers*, Snyder's (2006) discussion of Whewell's view on *causal* induction and Hacking's (2006) discussion of Mill's *co-instantiation* induction, in contrast.

[206] 'Mechanism' is used here in the simple/neutral sense of a series of factors spatio-temporally related, between which non-accidental relations are observed to obtain in experimental conditions, and which are correlated with statistical assessments from the level of populations.

numerous population-level assessments.[207] It is known of this relationship that due to a molecular mimicry between the TSH receptor from the level of thyroid membrane and certain proteins located in the retro-bulbar, pre-adipocyte soft tissue, TRAs (the antibodies directed mainly against the TSH-receptor of the thyroid membrane)[208] reach the orbital muscles and fibroblasts. Here, they induce an inflammatory infiltrate composed mainly of T-lymphocytes and macrophages. The cytokines produced by the latter stimulate the secretion of certain glycosaminoglycans (mainly hyaluronate and chondroitin sulfate) by the fibroblasts, with hydrophilic consequences.

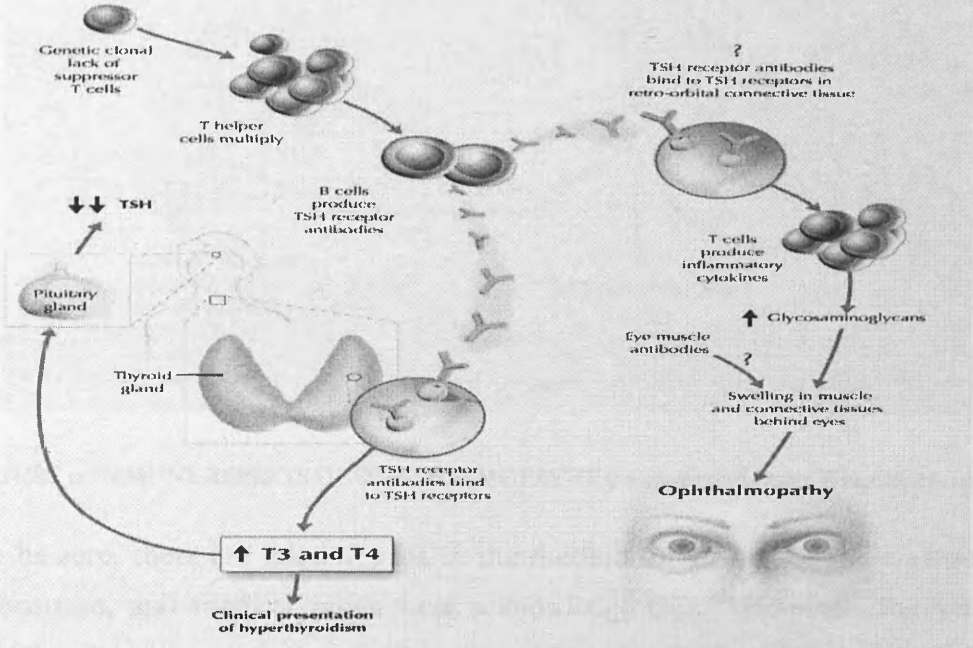


FIGURE 10 **OPHTHALMOPATHY MECHANISM (Reproduced from Ginsberg 2003)**

---

[207] See, e.g., [41] for *in vitro* laboratory assessments, [42] for animal testing evaluations, and [43] for epidemiological assessments.

[208] TSH (thyroid stimulating hormone) is a hormone secreted by the hypophyseal gland which controls the activity of the thyroid gland. Normally, TSH reaches a certain receptor located on the thyroidian membrane and enables a series of intracellular reactions in the thyroidian follicles that culminate in the synthesis of thyroxin (the main thyroid hormone) and triiodothyronine (thyroxine's active metabolite). When TSAs (the antibodies produced by lymphocytes as a result of 'perceiving' the TSH-thyroidian receptor as 'foreign', i.e., as an antigen) target the respective receptor, they enable the same set of intracellular reactions but in an exacerbated way. Informally put, this is the main reason for the emergence of Graves' disease symptoms. Thyroid normally controls the base metabolism rate. In Graves' disease, via the abovementioned determining properties located at the microcellular level, the metabolic influences of thyroid are exacerbated, and hence the superficial manifestations (the symptoms) show up.

The resulting oedema, together with the presence as such of the inflammatory infiltrate, leads to the protrusion of the eyes and the associated signs of exophthalmia—retrobulbar pain, pain on eye movement, eyelid erythema, conjunctival injection, chemosis, swelling of the caruncle, and eyelid edema.

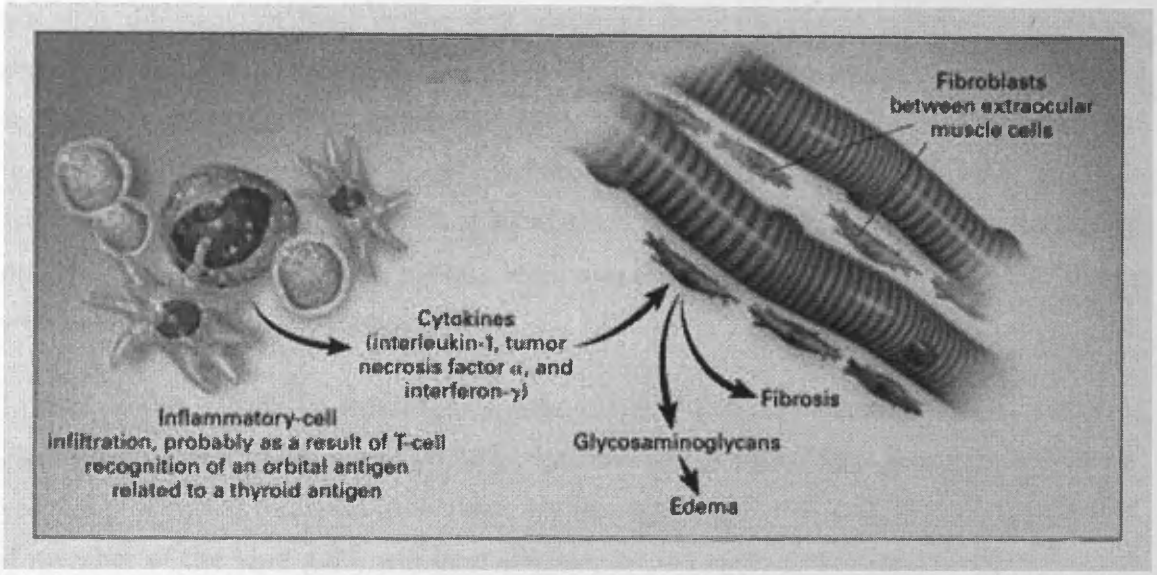


FIGURE 11 **IMMUNE ASPECTS OF OPHTHALMOPATHY (reproduced from Wheetman 2000)**

To be sure, there are missing gaps in the mechanism leading to the emergence of eye protrusion, and medical researchers acknowledge this.[209] However, the yet unknown details notwithstanding, by all of medicine's standards, the relation between TRAs and exophthalmia is non-accidental, as I have said.[210] That is why causal inferences can be framed for organisms with detected TRAs levels in blood vis-à-vis the presence of exophthalmia. It might be that these causal inferences, as all the other involving members of the GD kind, are not one hundred percent sure, as in the case of gold.[211] But they show that the GD kind fulfils criterion II, just like gold, without any ontological gap.

---

[209] See Wheetman (2000) for instance.

[210] A more detailed discussion of causal relations in medicine will be undertaken in chapter 4.

[211] Parenthetically, one could argue that the causal inductive inferences in which gold members are involved, as the inductive inferences across *all* sciences, are not universal. They always admit exceptions related to the presence or absence of certain other factors that enable or interfere with the manifestations of causal powers; see Cartwright 1989, 1998) and Lowe (1989) for a discussion directly framed in the (metaphysical) terms of causal powers.

3.1.2 *THE CO-INSTANTIATION INDUCTION REQUIREMENT (CRITERION V)*[212]

Criterion V says that, besides allowing causal inductions, inductively rich kinds should also enable us to formulate co-instantiation inferences. Such co-instantiation inferences, as discussed in chapter 1, are about the very possession of determining properties. They are thus different (at least in the first instance) from the causal inferences revolving around criterion II. Whereas the latter are based on the relation of causal dependence of superficial properties on the determining properties, the former concern the *co-instantiation* of determining properties themselves. In the notation I have adopted, criterion V says that, for any individual *a*, given a kind K characterised by a pattern of determining properties $[p_1, p_2, p_3....p_n]$, were *a* to be a member of K, *a* would, or probably would, co-instantiate $[p_1, p_2, p_3,....p_n]$.

In fact, without such co-instantiation inferences, the aforementioned use of kinds in causal inductions could not be possible. Informally, the rationale can be put as follows: yes, when it comes to causal inductions, we can expect that the next observed particular, if member of the kind gold, will be malleable, for instance, given that this kind has the determining property of having the valence shell electrons delocalised. But why should we expect the next member of the kind gold to have the valence shell electrons delocalised? More generally, why should we expect that the future observed kind members will present the same pattern of determining properties? Co-instantiation inferences are needed in order for the causal inferences involving superficial properties to make sense.[213]

Again, gold fulfills this requirement with a vengeance. The determining properties that characterise the gold element and its metallic complexes are so tightly united that philosophers and lay people alike often and in various contexts use the locution 'Z=79' to refer to the microstructure in question, even though the properties involved are more complex than a predicate referring to any atomic number may convey. The atoms

---

[212] I shall come back to criteria III and IV later. The reason why criterion V is discussed beforehand is that it is connected with the theme of induction and it has important argumentative links with criterion II.

[213] In other words, the inference 'Were *a* to be NK, *a* would, or probably would, instantiate $[q_1, q_2,....q_n]$' depends upon the inference 'Were *a* to be NK, *a* would, or probably would, instantiate $[p_1, p_2,....p_n]$'

composing gold samples have an electronegativity of 3.54 on the Pauling scale, an ionisation energy of 890.1 kJ/mol, a certain distribution of electrons in orbitals (including, as in all transitional elements, a partly filled d sub-shell (5d) that confers to it multiple oxidation states), 118 neutrons in the stable state, and varying numbers in its 18 radioactive states, not to mention the properties gold has in its metallic form.

Of course, that gold has an atomic number (Z) of 79 conveys a very important property of gold atoms and gold samples in general, and it has, in addition, a heuristic use in distinguishing the element at hand from all the other elements. But indeed, besides protons and their number, the determining properties present at the micro-structural level in question are more numerous. Now, why are the determining properties in gold's case thus united? And, more generally, what justifies, in the end, the co-instantiation inductions that can be formulated for the kinds satisfying criterion V?

Three different answers could be offered to such an inquiry. One is causal (and metaphysically quite uncomplicated) whereas the other two are non-causal and introduce intricate metaphysical notions – essences and substantive universals. One simply says that due to the *causal* structure of the world, the determining properties are agglutinated in certain patterns, which admit exceptions on the level of individuals. The other two say, in one way or another, that kinds *must* instantiate the *same* pattern of determining properties if they are to be natural at all.

It is interesting to note first that the causal answer in fact transforms the co-instantiation inferences into (more or less) causal inductions, thus bringing criterion V closer to criterion II. The causal answer explains that an individual *a*, if member of a kind K, would, or probably would, instantiate $[p_1, p_2, p_3....p_n]$ since *a*, as a member of K, participates at the level of its determining properties in causal interactions in which are involved various

internal and external mechanisms.[214]

These mechanisms could agglutinate [p_1, p_2, p_3....p_n] perfectly, such that *a* would instantiate all of the properties in question, or (depending upon circumstances) imperfectly, such that exceptions could show up at *a*'s level.

On the non-causal side, in contrast, the other two answers imply that the co-instantiation inferences are strictly separated from the causal inductions that the previously discussed criterion was concerned with. These two answers hypothesize that, if member of K, *a* should instantiate *all* of the properties of the pattern [p_1, p_2, p_3....p_n], either because K has an 'essence' and [p_1, p_2,...,p_n] are 'essential' properties,[215] or because *a* instantiates a substantive universal K which, metaphysically, imposes the instantiation of all of the [p_1, p_2, ...p_n] by *a*.[216]

In the present (sub-)section I shall look primarily at the first, causal option, which, as we shall see, is sufficient for the modest purposes of our comparison. I shall then comment rather briefly on the other two (non-causal) answers at the end of this section. I shall insist however upon the significance of metaphysical elements such as 'essences', 'essential properties' and 'substantive universals' in the following sections (when analysing criteria VI_a and VI_b).

The point I wish to make regarding the co-instantiation induction requirement is that the first, causal option does justice to gold *and* GD. In both cases, that is to say, we have a

---

[214] 'Mechanisms' are cited here simply in order to designate those internal and external causal elements that are involved in the agglutination of determining properties. I should point out that there are specific problems with the metaphysics of causation that involves the notion of mechanism (see for instance Woodward 2008, Salmon 1998 and Cartwright 1999).We do not need to go into these issues at this point though. To invoke mechanisms, in this context, is simply to make more intuitive how the 'causal structure of the world' intervenes into the co-instantiation of the 'profound' properties of kind members. Significantly, Boyd (1991, 1999), who introduced the causal story in the topic of natural kinds and the properties that characterise them, refers frequently to 'causal mechanisms', without discussing in any detail the ontological niceties associated with mechanism causation. It is worth noting that generally Boyd is a non-Humean about causation (Boyd, 1995: 365-7).

[215] *Cf.* Ellis (2002) Wilkerson (1995) Harré (2005)

[216] See Ellis (2001) Note that not all kind theorists believe that substantive universals induce the co-instantiation of all the properties of a characterising pattern on the level of particulars, as Ellis does. Lowe (2001) does not, even though he holds that on the level of universals, the patterns in question 'characterise' *in integrum* the substantive universals they correspond to; see Lowe (2006: 155, 169, 170, 173). See also Heil (2006) for a critical discussion of Lowe's position on (substantive or not) universals. I shall return frequently to Lowe's approach in the subsequent sections of this chapter.

causal story that explains the unity of the respective determining properties. In the case of gold, the story would appeal to a combination of classical elements of quantum mechanics, for instance, the nuclear and electromagnetic forces, Pauli's exclusion principle for quantum states, the Fermi energy, etc. That the determining properties of gold (including its number of protons) are *co*-instantiated is arguably due to causal interactions (such as the interactions governed by the colour forces holding gluons and quarks together and mediated by gluons themselves, the interactions between nucleons governed by the strong nuclear forces and mediated by the exchange of pions, the interactions between protons and electrons governed by Columb's force, etc).

Should any aspects of this picture appear as non-causal,[217] the causal story of the creation of (chemical) elements could also be invoked. Gold was one of the elements with an atomic number greater than 26 that was produced by the S process (slow-neutron-capture-process) in stars some 100 million years after the Big-Bang (when the first galaxies and stars were created) and it continues to be thus produced in stars (either by the S process, or the R process of rapid neutron capture in core-collapse supernovas).[218]

In the case of the GD kind, there is, on the one hand a mixed mechanism in which external environmental factors, such as smoking or iodine intake, and internal factors, such as various genetic predispositions, cause organisms to produce the above-mentioned TRAs—the antibodies directed against their own thyroids.[219] On the other hand, there are correlated internal mechanisms in which the over-stimulated thyroids produce excess amounts of their hormones, thereby augmenting the physiological effects, and finally creating a cluster of the above-mentioned 'pathological' determining properties (which

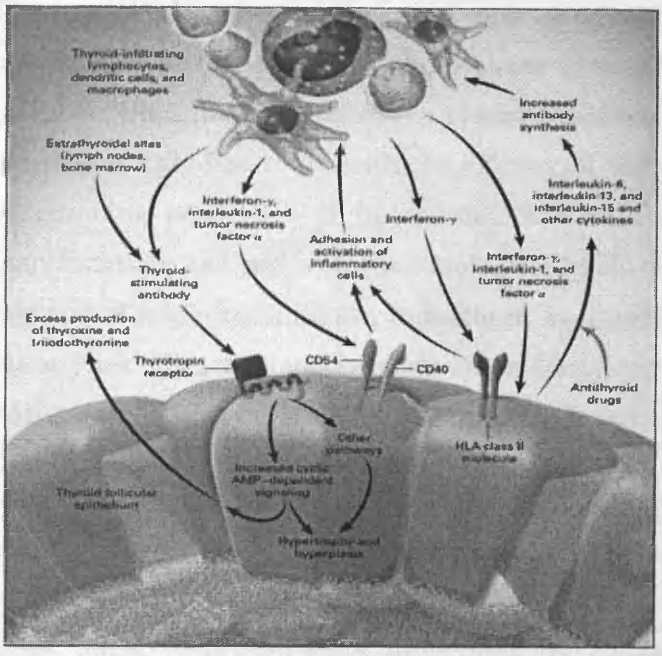
---

[217]One could argue that there are determining properties of gold that do not result from causal interactions *simpliciter*. For instance, the distribution of electrons in orbitals takes place according to Pauli's exclusion principle which simply forbids that two identical fermions may simultaneously occupy the same quantum state. Pauli's principle follows from the anti symmetry of the wavefunction(s) for fermions, which is commonly considered as being simply a principle of nature (see for instance, Fenyman 1966: section 4-4 and Pauli's own rendition in his 1946:17). A causal interpretation of Pauli's principle could be formulated in the framework of relativistic quantum field theory in terms of relativistic local causality or in Bohmian mechanics; *Cf.* Bohm *et al.* (1955) and Cushing (1993). For a philosophical discussion see van Fraassen (1980: 122-124), Lewis (1986: 222–223) and Salmon (1989: 159-164).

[218] See for example Wallerstein, G. *et. al.* (1997)

[219] See Holm (2005), Rapoport *et. al* (1998), Jacobson (2007), Manji (2006)

part of the cluster precisely, depending on the characteristics of each organism).[220]



Various immune cells (lymphocytes, dendritic cells, macrophages) initially coming from extra-thyroidal sites (lymph nodes, bone marrow) produce thyroid stimulating antibodies (TRAs) that attack a certain receptor of the thyroidian membrane (the thyrotropin receptor). The overstimulation of this receptor produces the hypertrophy and hyperplasia of thyroidian cells followed by the excessive production of hormones (thyroxine and triiodotironine). The immune cells synthesize also various cytokines (such as interferon g, interleukin 1) which amplify the inflammatory associated process and induce the expression by the thyroidan cells of receptors (CD 54, CD 40) and immune enhancing molecules (HLA II).

***FIGURE* 12 A PART OF THE CAUSAL MECHANISM INDUCING THE PRODUCTION IN EXCESS OF THYROID'S HORMONES. (Reproduced from Wheetman 2000)**

These 'causal stories' back up the use of co-instantiation inductions both for gold and the GD kind. Again, there are varying degrees of certainty regarding the ensuing inductions. In the gold case one can expect, say, that the next observed member *will* have the shell electrons delocalised in its metallic structure. In the case of the GD kind, one can expect, say, that the next observed member is to have, with a certain probability, mitochondria increased in number and size in a cells' cytoplasm. Alternatively, one can expect that it *will* have such mitochondria only if a host of certain other conditions are in place.[221]

However, just as in the case of the causal inferences involved in criterion II, one can argue that this difference in the certainty of co-instantiation inferences concerning determining properties does not mark out an ontological gap between the two kinds, from the point of

[220] To recall, thyroid is the gland that controls the base metabolism rate and hyperthyroidism mainly exacerbates these metabolic effects.
[221] See Brent (2008)

view of criterion V.

The causal elements in the gold case produce a 'perfect' pattern of determining properties. The causal elements in the case of the GD kind allow exceptions in that kind members instantiated parts of the cluster in question. The difference in the certainty of inductions only has to do with the efficacy of the causal paths underlying the unity of determining properties. It has to do, that is to say, with causation and its way of manifestation, and not with the ontological levels onto which the various kinds could be mapped. For co-instantiation inductions, it is sufficient that causation, as such, is in place. How efficacious mechanisms are in clustering properties is simply an issue that has nothing to do with criterion V and the metaphysics of kinds.

Now, I have mentioned in the beginning of this sub-section that besides the causal answer to the unity of determining properties, there are two alternative (and connected) solutions to the question of how members of a kind satisfy criterion V. They explain how kinds allow co-instantiation inductions by appealing to 'essences' and 'substantive universals'.[222] These are two venerable metaphysical notions, whose involvement in criteria VIb and VII has been explored in chapter2. I do not think that we should dwell upon them, in the context of criterion V (and II), for a very simple and yet convincing reason.

Insofar as criterion V (and II) is (are) concerned, essences and substantive universals can only play a redundant role. In fact, appealing to them is redundant on two (related) scores that concern criterion II and the very existence of these metaphysical elements, respectively. For one, calling upon 'essential' properties/essences of kinds, and/or the instantiation of substantive universals, cannot shed any (further) light on the capacity of kinds to participate in inductions. Once it is established on causal grounds that, for any individual *a*, if it is a member of the kind K then *a* will, or probably will, co-instantiate the pattern of properties $[p_1, p_2,..., p_n]$, further invoking the fact that *a* might also instantiate a substantive universal K (or that $[p_1, p_2,..., p_n]$ are 'essential' properties) is not any more informative about (the possibility of) the respective inference.

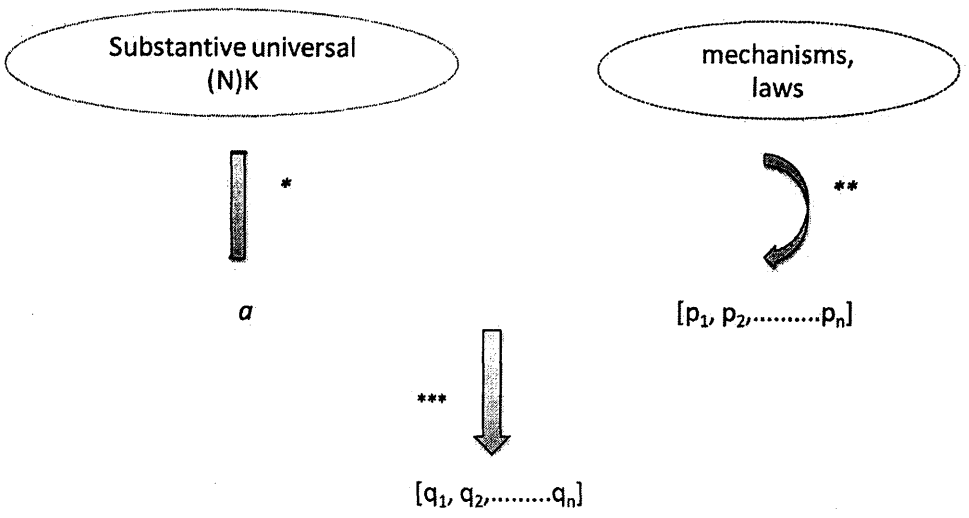
[222] See for instance Lowe (2001), (1989) for an appeal to substantive universals in order to explain why the kind properties co-instantiate in the way they do.

Of course, one might be interested in the certainty of the co-instantiation inductions, i.e. in whether they are universal or bear exceptions. Nevertheless, as I said, this aspect does not concern the metaphysics of (natural) kinds, from the point of view of criterion V – it can only indicate a difference of degree. Whether we have exceptions or not does indeed matter for kinds insofar as criteria III, VI_a and VI_b are concerned, and essences and substantive universals are actively involved in the discussion of criterion VI_b, as we shall see.

For another, if the causal response I have discussed above is coherent, then for reasons related to the principle of parsimony, the arguments in favour of the very existence of substantive universals or 'essences' of kinds should invoke at a metaphysical level *something other* than the need to justify induction for kinds.[223] There should be something else that warrants inflating our ontology with substantive universals or alternatively, with sets of properties constituting 'essences of kinds' - where the 'essential' properties are defined (a bit mysteriously) as those properties that all and only the members of kinds instantiate and must instantiate.

[223] *Contra* Lowe (2006: 135) who claims that substantive universals (and only they) could justify the co-instantiation of properties for natural kind members

FIGURE 13 **The co-instantiation of $[p_1, p_2, \ldots\ldots p_n]$ by *a* could be explained in causal way *simpliciter*, or by bringing up the instantiation of a substantive universals. Once the causal explanation is shown to be available though, the invoking of substantive universals becomes redundant.**



*** Between the individual *a* and the substantive universal NK there is a relation of instantiation, due to which a possesses the pattern $[p_1, p_2, \ldots\ldots p_n]$**

**** Alternatively, the possession by a of $[p_1, p_2, \ldots\ldots p_n]$ might be viewed as the result of the (efficient) causal structure of the world (mechanisms, causal laws)**

***** Between the 'profound' and the 'superficial' properties $[q_1, q_2, \ldots\ldots q_n]$ there is a causally efficient relation.**

I conclude that the co-instantiation requirement cannot situate gold and the GD kind on different ontological levels and that, more broadly, induction (be it of a causal or co-instantiation type) cannot reveal an ontological gap but a difference of degree. In the next section, the two kinds will be scrutinized from the point of view of an equally important criterion of natural kind membership – involvement in laws.

§3.2 THE REQUIREMENT OF LAW-INVOLVEMENT

A bolder development of the previously discussed requirement(s) says that natural kinds should be involved in (causal) laws of nature.[224] It might be for instance that electrons are involved in induction inferences of the type 'if all electrons observed so far exhibited this or that behaviour towards protons, then the next observed electron will, or will probably exhibit the same behaviour', or 'if something is an electron, then it will attract protons with this force'. However, one could point out that what is at work here is a *law* in which the electron kind is involved because *charge* is one of its characterising properties.

In other words, this criterion says that the properties shared by kind members should be not only natural and explanatively powerful but also 'nomic'. In my notation, it says that for an individual *a*—a member of a (candidate) natural kind NK in virtue of being characterised by a pattern of superficial properties $[q_1, q_2,..., q_n]$ and a pattern of determining properties $[p_1, p_2,..., p_n]$—the pattern $[p_1, p_2,..., p_n]$ instantiated by *a* should be involved in laws that govern/describe either the causal interactions of *a* whose effects are $[q_1, q_2,..., q_n]$ or the interactions of *a* with other individuals. Note that this requirement is different from causal induction requirement, insofar as kinds could participate in causal inductions without being involved in laws, and there are serious reasons to think this is the case.

For instance, special sciences represent for many scientific law theorists domains in which we do not have laws, at least in the canonical sense that characterizes the exact sciences.[225] Yet, insofar as one does not advocate reductionism (for instance of the biological or psychological to the physical level), natural sciences cannot be understood as lacking causation, i.e. as being characterized by arbitrary, accidental processes. Hence, in such domains it should be meaningful to discuss explanatory powerful properties and causal inductions even if the existence of nomic properties, in the full-blooded sense, is an undecided issue.[226]

---

[224] Criterion put forth by Bird (1998), Collier (1996), Ellis (2001) Cooper (2005) and Mackery (2005) among others.

[225] see Ellis (2001) and Carnap (1995).

[226] See for instance Boyd (1991), who persuasively discusses causal induction and causation in biology while accepting that the type of laws found in physics have no correspondent in the biological realm.

Now, the above formulation is not exactly appropriate, in the sense that it suggests a certain interpretation of what involvement in laws means for kinds.[227] The respective interpretation is not straightforward though, because of, broadly speaking, the various ways in which law statements can be framed. What I mean to say is that the danger of circular reasoning exists here to a certain degree. That is why, *inter alia,* the difference between laws and law statements will have to be considered in our discussion.[228] Importantly, the issue of what precisely laws represent is also involved here, as we shall see once the two kinds are set in contrast from the point of view of their nomic relationships.

What laws are gold and the GD kind involved in? One point that is obvious from the previous sections is that in the case of the GD kind we have exceptions on the level of the properties instantiated by kind members. Not all the superficial and determining properties of this kind are instantiated by all organisms with Graves' disease. We have, that is, a cluster of properties possessed by the GD kind members. In the case of gold, we do not have any exceptions because the various inductions that can be formulated for its members are universal.

This aspect is the cornerstone of our analysis from the point of view of the law-involvement requirement. A natural suggestion arises that the existence of exceptions betrays a difference in the laws in which the respective kinds are involved.[229] Simply put, an ontological gap could be seen to appear here – we seemingly have laws that do not admit exceptions in the case of gold and laws that do admit exceptions in the case of the GD kind (for instance, the law describing/governing the association between Graves' Disease and pre-tibial oedema).

However, we have two inter-related questions that need to be addressed at this point: are

---

[227] More precisely, it suggests that the truth-makers of laws are properties and that the main laws in which kinds are involved have a causal nature. Such an interpretation appears to set aside the fact that kinds might be the truth-makers of laws and that the laws in question might describe/govern co-instantiation relations between properties following from kinds' possessing an essence and/or representing substantive universals.

[228] I have purposely formulated the requirement in question in this form, in order to stress that on a *minimal* interpretation kinds are characterized patterns of properties (that is, at base kind members are simply individuals with similarities).

[229] See for example Millikan (1999: 47, 48, 54, 55) and Nickel (2008: 1, 15, 17)

laws that admit exceptions (ontologically) different from laws that do not admit exceptions? and, as I have already anticipated, what precisely is the nature of the involvement of kinds in laws? Both questions are in need of an answer if the criterion of law-involvement is to be reckoned with. In fact, the former question arguably depends on the latter, in the following sense.

If the involvement of (natural) kinds in laws is roughly viewed as being merely the way in which laws are supposed to govern/describe causal interactions in which kind members participate[230] - then the ontological discrepancy between laws that admit exceptions and those that do not needs to be elucidated. On the other hand, if the involvement of kinds in laws is investigated more carefully so that the distinction between laws and law statements (as well as the truth-makers of the latter) is heeded, then it might be that the ontological difference between exception-less and exception-ridden laws *does not have to be taken into account* in our discussion.

In fact, I consider the latter path more proper and wish to follow it. That is, I intend to look more closely at what involvement in laws is, in order to respond to the challenge posed by the causal induction requirement. Beforehand, I should say, however, a few words about the difference between exceptionless and exception-ridden laws, even if this vein of argumentation is, I believe, much less fruitful.

The existence of exceptions in a certain realm governed by a law could be viewed as resulting from that law being probabilistic or *ceteris paribus*. In other words, that the consequent of a law (statement) is not satisfied in states of affairs in which its antecedent obtains could be attributed to *ceteris paribus* clauses not being fulfilled or to the probabilistic form in which the consequent follows the antecedent. Notably, the *universally* quantified form of laws (statements) has been taken as canonical for the way in which the laws manifest themselves and such laws have been typically advocated for the exact sciences.

Exception-ridden laws have been taken to characterise the special sciences.[231] In fact,

---

[230] I take this as a translation of the equally rough dictum that 'laws "characterise" kinds'; see for instance Collier (1996: 2), Nelson (1990: 102) and Cooper (2005: 46)

[231] *Cf.* Psillos (2007: 38, 39,135, 136)

sciences such as medicine, sociology, economics, biology, etc. have taken on this denomination precisely due to the fact that their laws are supposed to admit exceptions (and hence to be *special*) in contrast to the 'normal' laws of the exact sciences.[232]

This denomination of course carries a certain qualitative burden (from an ontological point of view) in the sense that full-blown laws are supposed to be the ones from the exact sciences. It is exactly this burden that seems to incline the ontological balance towards gold as being 'more natural' than the GD kind.

Now, there are possible means of dispensing with this ontological burden, mainly by pointing out either that in the exact sciences probabilistic laws are present (at the level of quantum mechanics) or that the *ceteris paribus* clauses are needed for all sciences, across all domains, in order for their manifestations to show up on the level of particulars.[233] The existence of probabilistic laws in quantum mechanics could then be used in order to argue that if such laws show up at the most fundamental level, their ontological status could not be below the status of universal laws. On the other hand, the putative omnipresence of *ceteris paribus* clauses could arguably be employed to modify the canonical form in which laws have been traditionally construed.[234]

This vein of argumentation is not very fruitful, however. Exploring it would have to disentangle a host of very intricate issues. For instance, the probabilistic laws could actually be construed as exceptionless, universal laws if, on the level of their corresponding statements, the probabilistic clause is moved into the consequent.[235] On the other hand, the probabilistic form in which laws of the special sciences (medicine included) are sometimes expressed might not be genuine, law-like frequencies,[236] but might represent artefacts of the statistical means of analysing populations, reflecting our ignorance of certain other conditions that could transform the law conditionals in question into strict ones.[237]

---

[232] See Fodor (1974) and Millikan (1999)
[233] See for instance Mitchell (2000), Cartwright (1999) and also Lowe (1989)
[234] See Mitchell (2000)
[235] See Cartwright (1989:105). In fact, this is the classical form in which probabilistic laws are expressed (on a [relative] frequency interpretation of probability).
[236] See Cartwright (1989:142, 143)
[237] *Cf.* Hausman and Woodward (2004) *apud* Cartwright (2007b:107)

Were such conditions to be discovered, they would *prima facie* qualify as *ceteris paribus* clauses. But such clauses engender other problems of their own. For one, they (or at least some of them) could also be incorporated into the antecedents of law-statements (antecedents which would thus become very large) such that we end up having universal nomic statements.[238] For another, these clauses are subject to substantial criticisms claiming that, epistemically and semantically, they are nonsensical. [239]

All in all, no matter how the notion of law is stretched and twisted, the existence of exceptions still seems to indicate for our comparison an important difference in laws that is reflected (or originates in) the nature of kinds. Take the case of *ceteris paribus* laws. Notwithstanding the discussion over the intelligibility or unintelligibility of *ceteris paribus* clauses, over how they could be inserted into the antecedents of laws, etc. one can still argue that since laws 'characterise' kinds, the existence of exceptions in the manifestation of laws betrays a difference in kinds. In our specific circumstances, one could still hold that the *ceteris paribus* conditions of the laws governing the emergence of ophthalmopathy, for instance, arise because the members of the Graves' Disease kind are not exactly similar. Since, in contrast, the members of the gold kind are exactly similar and the laws which 'characterise' them do not admit exceptions, the two kinds still appear to be situated on different ontological levels, from the point of view of the law-involvement criterion.

We have, however, another vein of argumentation at our disposal, which is more powerful because it looks directly at what the involvement in laws means for kinds. In this vein, the difference between laws and law-statements, as well as the truth-makers of the latter, are crucial. These aspects could be employed in order to argue that the existence of exceptions does not matter with respect to the metaphysics of kinds, at least as far as criterion of law involvement is concerned.

Let us take a (very) schematic scenario in which a kind K is characterised, say, by the properties q_1, q_2, q_4 and p_3, p_5, p_7. The properties at hand are involved in causal

---

[238]See Cartwright (2002b)

[239]See Earman *et. al* (2002) and also Drewery (2001), Woodward (2002) for discussion. A reply to Earman *et al.*'s (critical) approach to *ceteris paribus* clauses can be found in Cartwright (2002b); Lange defends the meaningfulness of these clauses from a different perspective in Lange (2002)

interactions, and these causal interactions are in turn governed/described by laws. When it comes to the involvement of K in the laws in question (i.e. when it comes to how laws 'characterise' K), one could choose various nomic statements that directly connect K with them. For instance, it might be said that *it is a law for the members of the kind K* to instantiate q_1 and q_2, once they instantiate p_5. If the determining properties are under focus, one might say that *it is a law for the members of the kind K* to instantiate p_5, once they are engaged with such and such external and internal mechanisms. If *a*, as a member of K, does not instantiate, say, q_2, or p_5, then, as I said, K can appear to be on an inferior ontological level in comparison to another kind K* whose members instantiate all of its characterising properties, since the laws of K appear as *ceteris paribus* (or probabilistic) whereas the laws of K* are strict.

Nonetheless, as I said, the question worth addressing is: irrespective of how nomic statements are formulated, can one really transfer the ontological consequences (if any) of exceptions from the level of laws to the metaphysical level of kinds? My conjecture is that they cannot be transferred if kinds are not the truth-makers of these statements. In other words, if all there is in the involvement of kinds in laws is that certain law-like statements can be put to use for various practical reasons,[240] then the existence of exceptions is ontologically innocuous for kinds.

My argument appeals in the first instance to the causal picture framed in the previous section concerned with induction. We have causal interactions that result in the agglutination of properties. Causal mechanisms, external and internal, produce the co-instantiation of determining properties. The latter in turn cause the emergence of superficial properties. Now, admittedly, laws govern these causal interactions.[241] More precisely, the *efficacy* of mechanisms is underlined by causal laws, such that whether the determining properties are perfectly agglutinated (i.e. exceptionless vis-à-vis each kind member) or imperfectly agglutinated (i.e. leading to exceptions for some or all kind members) is due to laws manifesting themselves in different ways (or, we could admit, to laws having a different nature). In turn, laws with potentially different natures similarly

[240] Indeed, it could be argued that in scientific practice such law statements mentioning kinds are indispensable or at least very helpful - see the case of biology for instance.

[241] If we leave aside positions such as Mumford's from his (2004) according to which laws are metaphysically redundant (in an anti-Humean framework of causation).

underline the efficacy of determining properties in producing the superficial properties.[242]

However, how efficient mechanisms (and the determining properties) are in producing their effects is not significant for kinds from the point of view of induction. This simply represents a re-statement of the point argued for in the previous two sections. Similarly, I want to claim that insofar as laws are under focus, the efficiency of the causal paths travelling from mechanisms to determining properties and then to superficial ones could mark out a significant difference only if kinds were directly responsible, from a metaphysical point of view, for the efficiency in question.

To return to a previous line of argumentation, in what sense specifically could the *ceteris paribus* clauses attached to certain laws arise *because* the kind members are not exactly similar? In this interrogation, the precise sense of 'because' needs to be clearly unveiled. Does it betray the metaphysical 'responsibility' of kinds for the specific laws at issue? If the truth-makers of laws are not kinds, then what we have here is a perfect illustration (in the realm of metaphysical reasoning) of the fallacy that Reichenbach's principle of common cause seeks to avoid (in the realm of causation).[243]

This principle says that for any correlation between two factors, one causes another only if there is not any other factor that, if considered, would screen off the initial correlation. By analogy, we always have *ceteris paribus* clauses attached to law statements for kinds whose members are not exactly similar, i.e. we always have a 'correlation' between kinds whose members are not exactly similar and *ceteris paribus* clauses attached to law-statements. Nevertheless, this 'correlation' should be metaphysically significant on the level of (natural) kinds only if kinds were the truth-makers of such law-statements. If kinds do not play the respective role of truth makers, then the metaphysical 'correlation' between the members of a certain kind not being perfectly similar on the one hand, and the presence of *ceteris paribus* clauses on the level of law statements, on the other hand, must simply be due to another common metaphysical 'cause' – whatever the truth-maker

[242] See Bird (2007)
[243] For a discussion of Reichenbach's principle, see Cartwright (1989)

of the law-statements might be.[244]

To put it differently, my point is that metaphysically, the level of laws and the level of kinds are separate unless kinds are the truth-makers of law statements. The bearing of the *ceteris paribus* clauses-related or probabilistic exceptions on the nature of the causal laws in question can be transferred to kinds and can thus amount to an ontological difference *only* if kinds are taken to make law-statements true. This is just to say that the requirement of law involvement - *where* laws are attributed some sort of 'canonical' form *and* their manifestations are deemed to be exceptionless - is justified only if kinds are the truth-makers of laws. Otherwise, the nature of laws is separated from the level of kinds.[245]

There is indeed an alternative metaphysical approach in which kinds do represent the truth-makers of laws – just in case one accepts 'essences' of (natural) kinds and/or construes (natural) kinds as substantive universals. This view is not very orthodox in contemporary philosophy of science, it should be said, and I have already argued that the need to warrant induction cannot be a reason to introduce such elements into one's metaphysics of kinds.[246]

I conclude that the criterion of law involvement cannot indicate an ontological gap between the two kinds under comparison.

---

[244] Be it represented by brute regularities (*a la* Hume) or refined ones (*a la* Lewis 1986), generalisations invariant to manipulation and useful for prediction (*a la* Woodward 2002 or Mitchell 2000), capacities possessed by individuals (*a la* Cartwright 1989), single track potencies identified with properties (*a la* Bird 2007), clusters of powers associated with properties (*a la* Mumford 2004), or whatnot. I need not delve into this issue here.

[245] I would not want to commit myself to the stronger claim that the requirement of law involvement, *as such*, is unjustified. I we construe laws in a broad sense, as being capable of manifesting themselves with exceptions on the level of particulars, then it might well be the case that all natural kinds are involved in laws. Take the particular discussion of clusters and kinds. As we will see in the next section, one can plausibly view clusters as the result of 'the causal structure of the world'. But if kinds are construed as clusters and, further, causation is construed as arising due to the existence of laws, then it might well be that laws are involved with all cluster-kinds. Nevertheless, no problem should arise in this case for our comparison, vis-à-vis criterion IV. That is precisely because some clusters are perfectly 'agglutinated' while others are less perfectly 'agglutinated', due to the different levels of efficiency for causation. If laws are taken to underlie *all* causation phenomena, then laws should also be taken to admit exceptions on the level of particulars (for the special sciences domains).

[246] See also Salmon (1982) and Mumford (2005) for a devastating critique of the essentialist approach to natural kinds, either in its Kripke/Putnam type (plain semantic form) or in the Ellis type (metaphysical form). Substantive universals are convincingly shown as redundant for the metaphysics of laws in Bird (2007)

§3.3 THE CLUSTER REQUIREMENT (CRITERION VI_A)

From my previous discussion about laws and induction, it is clear that the existence of exceptions on the level of properties instantiated by kind members is an important aspect of my comparison, which needs to be considered when approaching almost all the criteria of membership. There is however, a criterion explicitly concerned with the fact that the members of a kind might not be exactly similar.

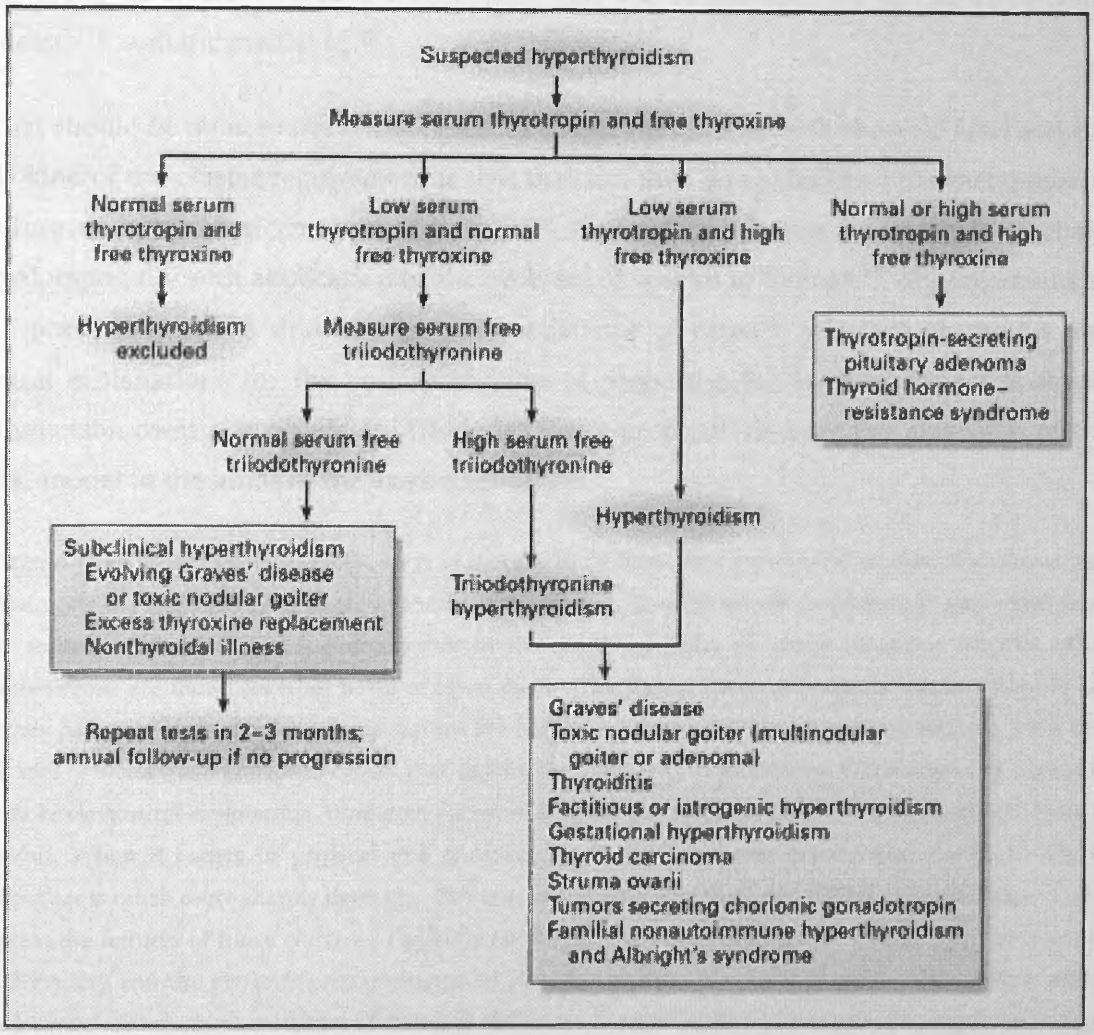
Criterion VI_a (the cluster requirement) discusses what limits the patterns of determining properties should have and simply says, in the notation I have adopted, that the properties $[p_1, p_2,...p_n]$ characterising a (candidate) natural kind K should form a cluster as a result of the causal structure of the world. Neither all of $[p_1, p_2,...p_n]$ nor any part of them need to be both necessary and sufficient for membership. All we need is to follow the patterns in which the determining properties are co-instantiated. We might find some 'divisions' that are fuzzy (and criss-crossing), and other 'divisions' that are more finessed, as it were. They all fulfil the cluster requirement though, insofar as the corresponding regions of density in the logical space of the determining (and superficial, one could add) properties in which the kind members can be localised, have a causal explanatory background.

It is obvious that both gold and the GD kind fulfil criterion VIa. In the case of gold, we even have necessary and sufficient conditions for membership. In the case of the GD kind (like in many other diseases and disease kinds), we only have certain necessary conditions.[247] More precisely, we have on the one hand a condition for the establishing of hyperthyroidism, namely low serum levels of thyrotropin and high levels of free thyroxine and/or triiodothyronine(see figure below), [248] and on the other hand the condition that TRA_s should be detected in the blood, which singles out Graves' Disease among other

[247] Other conditions with necessary (but not sufficient) conditions are for instance glomerulonephritis, hepatic cirrhosis, anemia and leukemia. See Lote (2000), Burgun *et. al.* (2005).
[248] In the earliest stages of Graves' disease, patients may have only increased secretion of triiodothyronine. *Cf.* Wheetman (2000)

hyperthyroidian states.[249]

FIGURE 14 DIAGNOSTIC ALGORITHM (REPRODUCED FROM WHEETMAN 2000)



Other diseases in somatic medicine have sufficient but not necessary conditions and arguably, there are also illnesses in which necessary and sufficient conditions are in

[249] In practice, the examination establishing the existence of TRAs is not always performed because the presence of symptoms and signs like goiter and exophthalmia is taken as sufficient for diagnosis *Cf.* Wheetman (2000). TRAs are considered worth investigating though in doubtful cases (along with an antibody directed against the thyroid peroxidase) with suggestive epidemiological background (for instance, for patients coming from iodine-excess areas) or for evaluative roles in treatment; see Davies (1998) for discussion.

place.[250] Insofar as the clusters of determining properties for each disease are non-accidental however, i.e. are the result of the causal structure of the world, criterion VIII is also satisfied for them. Indeed, with the exception of idiopathic diseases, in which (at the very least) an epistemic obstacle is present, this can be affirmed for all life-threatening illnesses in somatic medicine.[251]

What should be pointed out with respect to the satisfaction by both the gold kind and the GD kind of the cluster requirement is that they are thus integrated into the metaphysical picture of homeostatic property kinds (HPC) introduced in the literature by Richard Boyd, especially with application to the problem of species in biology.[252] My arguments in the previous sections drawing upon the existence of causally efficient properties and causal explanations for the co-instantiation of properties for *both* of the kinds under comparison owes also a debt to Alexander Bird's proposal vis-à-vis the extension of the HPC model to the kinds in the exact sciences.

'Richard Boyd takes biological kinds, such as species, to be homeostatic property clusters. The idea is that certain sets of properties tend to group themselves together. Thus, given all the biological properties there are, some combinations are found together in the same particular on many occasions whereas other combinations are found together never at all or rarely. The logical space of property combinations is not equally occupied by particulars. Some regions are highly populated whereas others are empty.... Thus it is *the laws of biology and biological causes* that explain the clustering of properties. *The existence of biological kinds has a natural explanation, ultimately in terms of laws. The same may well be true of natural kinds in general.* When it comes to physical and chemical kinds, the laws may ensure that the clustering of properties is much more sharply defined.... *The laws will explain why there are certain clusters*; they will also explain the *natures* of those *clusters*—the *loose and vague* clusters in biology, the partially precise clusters of chemistry and the *perfectly precise* clusters of particle physics... The laws of nature will explain why—necessarily—there are no members of chemical and microphysical kinds that lack certain properties, why of

[250] Sufficient but not necessary conditions are met mostly in rheumatoid diseases such as lupus erythematosus systemic, rheumatoid arthritis, etc.

[251] Idiopathic pathologies are those illnesses whose causes are unknown (or partly unknown) in contemporary medicine. In other words, they simply represent sets of symptoms observed to co-occur, without a well-delineated causal base. Parenthetically, this does not mean though that the diseases in question are 'rag-bag' conditions (to use Cooper's expression), since they can efficiently be treated. For example a considerable proportion of hypertension cases have an unknown basis and yet, the general anti-hypertensive treatment is efficacious in these cases as well. See Kasper *et. all* (2008)

[252] See Boyd (1991). Clusters, as such, have been introduced into the biological discussion by Dupré, without the attending causal story (Dupré, 1981). This approach, which does not mention the causal aspects, limits much the appeal that clusters have in the kinds discussion. Some authors prefer Dupré's version rather than Boyd's one, though; see for instance Cooper (2005).

necessity certain properties cluster together in a partially or fully precise manner.' (Bird, 2007: 209-210, italics added)

The (more or less perfect) agglutination of properties as a result of various causal interactions particulars are engaged in, both in the exact and special science(s), is a central feature of Bird's proposal. Bird takes the causal interactions to be underlined by laws and laws in turn are taken to express in an anti-Humean vein the (metaphysically necessary) association of causal powers to certain properties.[253]

Notably, since properties are then the truth-makers of law-statements (or of nomic subjunctive conditionals describing laws' manifestation), we can infer that the problems posed by criterion III (the law involvement requirement) do not show up in this model.

The loose and vague clusters of biology, suggests Bird, are to be explained by the biological laws that allow exceptions, whereas the perfect clusters of physics should emerge due to the exceptionless physical laws. In Bird's construal of laws as a metaphysical category, natural kinds, substantive universals and any sort of kind essences do not play any significant role.[254]

On a different note, it should be stressed that that almost all theorists of medicine who claim that medicine (and the special sciences in general) possesses a structure of natural kinds, embrace the idea of 'cluster kinds' as a central feature of their views.[255] I have noted in chapter 1 that, in a sense, my overall enquiry - which does not take up a certain conception of what natural kinds are, but deals with various criteria of membership – is simply meant to complement the 'classical' argumentation of these theorists intending to demonstrate that medical kinds fit coherently into a picture of 'cluster' kinds.

We have seen that the GD kind, just like gold, fulfils criterion VI_a. There are, however, two other criteria that make stronger demands about the limits of the patterns of properties instantiated by a (candidate) natural kind. I will look at them in the next

[253] Bird (2005)

[254] Even if, as we shall see in chapters 4, I consider the anti-Humean account of causation particularly suitable for medicine and medical explanations, I have tried in the part of the discussion directly concerning natural kinds to remain neutral vis-à-vis the Humean/anti-Humean variants of causal framework.

[255] See for instance Cooper (2005), Reznek (1987). The main views of these authors have been discussed in chapter 1.

section.

§3.4 THE N&S AND NON-OVERLAPPING REQUIREMENTS (CRITERIA VI_B AND VII)

Two other important criteria often called to pick out natural kinds are that their members should have necessary and sufficient conditions of membership (criterion VI_b) and that different natural kinds should not share members except when they are in a species genus hierarchy (criterion VII). In other words, from the 'divisions in nature' corresponding to natural kinds it is often asked to be clearly delineated (non-fuzzy) and non-overlapping (except in a genus-species hierarchy).

In the notation I have used so far, criterion VI_b says that, for any individual *a*, member of a kind K, the determining properties part of the pattern $[p_1, p_2,...,p_n]$ should be necessary and sufficient for membership. Criterion VII requires that, were *a* to also be a member of a different kind K*, then either K and K* are identical, or they are situated in a species-genus hierarchy. How does our comparison fare when these two criteria are reckoned with?

In the case of the gold kind, since its members possess as determining properties precisely the characteristics of the gold chemical element, *it is the case* that non-fuzziness and non-overlapping are met. That is, since the gold kind members could not possess only 'a part' of the determining properties found on the level of the respective chemical element, it is the case that the gold kind members fulfil the N&S requirement. In addition, given that there is discreetness (non-overlapping) between the gold element and all the other chemical elements, it is often argued that the gold kind members could not belong simultaneously to other kinds at the same (hierarchical) level.[256]

In contrast, GD is an overlapping, fuzzy kind, both at the level of symptoms and, more

[256] However, even if there is no overlapping on the level of chemical elements, as such, we have higher – level chemical kinds that could overlap; for instance, gold could be seen as part of a series of criss-crossing kinds such as metal, transitional element, substances whose crystals are face-centered cubic lattices, etc. Theorists like Ellis for instance, who backs up the non-overlapping requirement, insist on the discreetness from the level of chemical elements (Ellis 2001) and I shall assume, for the sake of the discussion, that gold does satisfy criterion VII.

importantly, at the level of 'pathological' determining properties. For one thing, obviously, an organism possessing the characteristics of Graves' Disease can simultaneously possess the characteristics of other diseases that are not in a species-genus relationship with Graves' Disease. Hence, criterion VII is not satisfied. For another, in order to be a member of the GD kind, an organism just needs to present a part of a cluster of 'pathological' determining properties.[257] As detailed in the previous section, we have some necessary conditions for membership, namely the possession of TRA_s antibodies, low levels of thyrotropin (TSH) and high levels of free thyroxine and/or triiodothyronine. No necessary and sufficient conditions are in place though.

I have formulated gold's situation vis-à-vis criteria VI_b and VII by saying that *it is the case* that the respective kind (as many other kinds in the exact sciences) is a non-fuzzy, non-overlapping kind. I wanted to emphasize that the gold kind corresponds to a certain type of division of nature, with certain characteristics. In chapter 2 I have argued, however, that kinds corresponding to 'divisions of nature' that do not present these characteristics should not be rejected as non-natural. That is because criteria VIb and VII can only be justified if the identity sense of carving nature is adopted.[258]What follows is that, in the context of the non-identity sense of carving nature, the N&S and non-overlapping requirements do not have to be fulfilled by the GD kind and the fact that the gold kind fulfils them does not have any special ontological importance.

To recall, both the Lockean strategy and the argumentation based on substantive universals cannot explain why the cluster kinds - which stand for divisions of nature that can be well-delineated but can *also* be fuzzy and overlapping - should not be considered

---

[257] See §3.1.2 in which some of the 'pathological' determining properties in question are summarised.

[258] As regards criterion VIb, for any individual *a*, member of a kind K, that the determining properties part of the pattern $[p_1, p_2,...,p_n]$ should be necessary and sufficient for membership only if the identity of a depends on all of the $[p_1, p_2, ...p_n]$ properties. If it did depend, then a could not lose p_3 and still remain a member of the kind K, for the simple reason that after the change, we would have instead of a another individual. As regards the hierarchy requirement, again, it would be justified only if the identity of (natural) kind members depended upon their membership. Let there be two kinds K and K* characterised by the patterns $[p_1,p_2,....p_n]$ and $[p_2, p_3....p_n, p_{n+1}]$ respectively, which overlapped (without being in a hierarchy), in the sense that an individual *a* instantiating $[p_1,p_2,....p_n, p_{n+1}]$ was a member of both. Suppose that *a* lost the property p_{n+1} in a process characterised by spatio-temporal continuity, continuity of causal history, etc. From the point of view of its membership in K, the individual resulting from the process should be (numerically) identical with *a*. From the point of view of its membership in K*, the individual in question should be different from *a*, since the kind membership (one of the necessary conditions for identity preservation) was lost. We would thus have a violation of the *tertium non datur* principle.

natural kinds. Take the 'essential' properties involved in the Lockean strategy. The properties characterising cluster kinds are microstructural, explanatory powerful properties, just as the properties characterising non-overlapping kinds with necessary and sufficient conditions of membership are. What could be the difference between the former type of properties and the latter?

As far as the notion of the essential is concerned, we have two options. Either there is no difference, if for a property to be *called* essential is to be micro-structural and explanatory powerful – in which case cluster kinds should be acknowledged as *natural* as well – or we have a difference, if for a property to be *called* essential is to be part of necessary and sufficient conditions of kind membership. Reckoning with this difference could not force us, however, to exclude cluster kinds and the fuzzy, criss-crossing divisions of nature from the *natural* kinds category because it simply amounts to affirming that some kinds (do not overlap each other and) have necessary and sufficient conditions of membership whereas others have more loose conditions of membership. To believe it amounts to something more is to decide *de dicto* that the fuzzy (and overlapping) divisions of nature are ontologically inferior to the non-fuzzy ones, i.e. to stipulate that the essential properties are necessary and sufficient for *natural* kind membership.

It might be that the divisions that are non-fuzzy in the actual world are metaphysically necessary so, i.e. are still non-fuzzy in all the possible worlds, perhaps as the result of causal interactions that agglutinate properties with more than physical necessity. Nonetheless, this aspect has to do with causation and other related issues - it does not tell us anything about the metaphysical category of natural kind and the different sorts of divisions of nature that should be acknowledged under its scope. The properties characteristic of kinds that are *called* essential lack a certain additional *de re* ingredient. As Stephen Mumford puts it "there should be something else, something else about the properties characteristic of kinds that makes them 'essential' [for *natural* kinds]."[259] This additional 'de re' ingredient, I have argued in chapter 2, cannot be represented by substantive universals.

Take as an example Harré's view on kinds, also discussed in chapter 2. Harré defines

[259] Mumford (2005: 424)

essences and 'essential' properties as follows.

"ESSENCES: A REMINDER 1. Properties of kinds a. Proprium: properties found in or displayed by every instance of a kind. b. Essence: properties necessary for a being to be a member of a kind c. Accident: properties found in some instances of a kind. Accidental properties displayed by some individuals in a group cannot be part of the essence of the kind. [...] Essences are sets of properties selected according to various criteria from the propria. In philosophical writings, 'essences' have four main features: i. They are immutable, the essence of a kind cannot change. ii. They are indivisible, a subset of the constituent defining properties of an essences is not an essence. iii. They are necessary, unless a being displays the properties defining the essence it is not a member of the kind. iv. They are infinite, that is an indefinite number of actual beings may realize an essence.... The real essence of type or kind is the cluster of properties necessary and sufficient for a being to be an instance of a type or a kind" (Harré, 2005:10, 11)

But what are the 'essential' properties, beside the fact of their possession being necessary and sufficient for kind membership? Well, it turns out that they are the intrinsic, structural properties that can be indicated as causative factors for the observable properties and behaviour of their bearers. But this is precisely the definition of determining properties we have dealt with in §1.2

"In practice, the properties comprising the real essence of a kind are often unobservables....Descriptions of the real essences of material stuffs are not 'definitions'. They are theoretically derived hypotheses *as to the constitutions or natures of the beings in question*. For example, that the hydrogen molecule is a diatomic structure of atoms consisting of a single proton and an orbiting electron, is a hypothesis as to the real essence or true nature of the light, inflammable gas that is prepared by the use of certain procedures. While real essences are not revisable, statements about real essences as empirical hypotheses are revisable and have often been revised." (Harré, 2005: 12, italics added)

The properties of the hydrogen molecule explain the nature of the light, inflammable gas we observe. Similarly, one could easily add, the atomic properties of the gold element and its metallic combinations explain the behaviour and properties of particulars that are yellow, malleable, fusible, soluble in *aqua regia*, etc. And similarly, the biological microstructural properties of the GD kind organisms explain their superficial properties and behaviours.[260]

[260] Of course, I must admit that the properties of gold explain almost *everything* that gold kind members exhibit, while the GD kind determining properties explain only a (small) part of how human organisms behave and what these organisms instantiate 'superficially'. It would be very easy to argue though that this does not mark out an ontological gap but indicates an (important) difference of degree.

Of course, one could even argue that, if the 'essential' properties possessed by the members of the gold kind are the ones that are causally efficacious for 'superficial' properties and behaviours, then both gold and the GD kind possess the corresponding explanatory/causally efficient 'essences'. But the bottom-line we are interested in is that such 'essential' properties, construed in Lockean tradition, cannot justify the N&S and non-overlapping requirements.

§3.5 THE NON-PHASE REQUIREMENT (CRITERION VIII)

Criterion VIII asks that the 'divisions in nature' corresponding to natural kinds should be non-phase (non-implicated in transitory transformations). In the terminology I have employed so far, it says that the pattern of determining properties $[p_1, p_2,...p_n]$ of a (candidate) natural kind K should not represent the phase of a more fundamental pattern of properties.[261] When kinds are involved in processes that have to do with their determining properties, the possibility that they might turn out to be phase kinds should always be considered.

The non-phase requirement has a straightforward importance for the case of pathological kinds, which are, of course, kinds of organisms and are also supposed to correspond to the classifications in terms of diseases that such organisms allow. To be sure, diseases involve numerous processes and, just to make matters worse, pathological conditions can be interpreted as being *fundamentally* processes. By focusing then on organisms that present certain symptoms and using the 'static' framework of natural kinds of *objects*, one appears to lose sight of the basic 'transitory' nature of diseases, as such. Pathological

[261] This is admittedly a non-illuminating definition, since what a phase represents is not elucidated. Recall though that the terminology I have employed was meant to help the exposition while preserving the content of the claims usually made about the metaphysical category of kinds in the literature. In the case of criterion VIII, the rather primitive role of 'phase' in its formulation reflects the ambiguity present in the literature regarding the ontological delimitation of phases. In fact, it could be said that the ambiguity stems not from the lack of a minimal indication of what a phase is, but from a plethora of characterisations of it that differ from author to author. Around this issue (i.e. the exact delineation of the phase/non-phase distinction) will actually revolve my argument in the present section.

kinds of organisms, that is to say, do appear as phase kinds.[262]

In the context of our comparison, it could be claimed that Graves' Disease is a process or a sum of processes undergone by human organisms. These processes alter organisms from having certain dimensions of eyes to an exophthalmic state, from having a certain intestinal transit rhythm to an accelerated one, from having a certain sugar and sodium cellular permeability to an increased one and so forth. The GD kind, it would seem, fails criterion VIII. In contrast, gold seems to be ideally suited to the framework of natural kinds of objects. The gold kind does not *prima facie* pick out an intermediary stage in a process undergone by individuals and thus it does appear to fulfil the non-phase requirement.

Are natural kinds of *objects* the wrong ontological framework for Graves' Disease? Or, in other words, is the GD kind a phase kind? In chapter 2 I have approached criterion VIII from a general metaphysical position, arguing that the 'Heraclitan' challenge that the special science kinds pose over the common-sense phase/non-phase distinction can be refuted only by adopting the identity sense of carving nature. In the present section, I will complement the respective line with a more focused and applied analysis for the two kinds under comparison. My response will not involve a definitive answer to the question whether the GD kind a phase kind. I shall rather be concerned, as in all the previous sections, with the possible ontological gap between the GD kind and gold. The conclusion I shall try to justify will be that, if GD is a phase kind, then gold is a phase kind as well. Conversely, if gold is considered a non-phase kind, the GD kind should also be construed in this way. The difference between the two kinds vis-à-vis criterion VIII, I will show, is in fact a difference of degree. I will also come back to the proviso regarding the identity of kind members (and the identity sense of carving nature at its joints) will also be maintained in this section, in a sense that will be made clear shortly.

Before approaching the main discussion, a brief clarification should be put in place. Criterion VIII has to do with the phase or non-phase nature of the patterns of *determining* properties, as such. Accordingly, the sort of processes that involve the superficial properties of kind members are not significant for the threat of representing a phase kind.

[262] Of all the kind theorists discussing the medical kinds, none of them pays any serious attention to the phase threat or to the fact that diseases might fundamentally be processes.

For a kind K characterised by the sets of properties $[p_1, p_2,...p_n]$ and $[q_1, q_2,...q_n]$, what criterion VIII is concerned with is whether the pattern $[p_1, p_2,...p_n]$ represents a phase in a process involving a more fundamental pattern of determining properties. The superficial properties $[q_1, q_2,...q_n]$ and the processes leading from $[p_1, p_2,...p_n]$ to $[q_1, q_2,...q_n]$ do not pose any threat from this point of view. These processes simply correspond to the causal interactions that should occur in any (candidate) natural kind characterised by determining properties.[263]

This amounts to saying that, when trying to dissipate the worry that diseases are fundamentally processes and that natural kinds of objects represent the wrong metaphysical framework for them, the sort of changes that lead to the emergence of the superficial, symptomatic properties starting from the determining 'pathological' ones, should not be a taken as a problem.[264] What we should worry about (from the point of view of the phase/non-phase distinction) are the processes due to which organisms come to possess or lose those determining properties as such.

For instance, in the case of the GD kind we have on the one hand the process(es) due to which organisms with high levels of thyroxin come to suffer from weight loss and accelerated intestinal transit. On the other hand, we have the process involved in organisms' acquiring these levels of thyroxin as such, i.e. the process due to which the TRA_s are created and directed against (a certain receptor of) the thyroidian membrane, with subsequent excitatory effects.

As regards the former type of (pathological) processes, my point is that acknowledging and paying full heed to their existence is not at odds at all with the hypothetical existence of natural kinds of (diseased) organisms. To the contrary, such kinds of organisms could help explaining the processes at hand and integrate them into a more complete ontological picture. Just as in the exact sciences to every (natural kind of) process there should in principle correspond a natural kind of objects (to use Ellis's framework, for

---

[263] *Cf.* Ellis (2001: 22,23,160,163)

[264] As it should be clear, I intend to argue that natural kinds of objects do not represent the wrong ontological framework for Graves' Disease, in the context of our comparison. This is not to say that I thus exclude the possibility of coherently construing diseases as (natural kinds of) processes. The latter is simply a vein of argumentation I do not pursue.

instance),[265] there does not seem to be any serious reason why in medicine a natural kind of organisms should not correspond in principle to every pathological process of this sort.

The rationale would be the same – given that pathological processes stem from the existence of causal interactions in which certain determining properties are involved, the possession of those determining properties can be used to circumscribe (candidate) natural kinds of pathological organisms. In our case, to the process that leads, say, from the high level of the Na^+/K^+ ATP_{ase} to the emergence of heat intolerance, there could very well correspond a (natural) kind of organisms, GD, whose members possess, *inter alia,* such levels of the ATP_{ase}. In a similar fashion, in the case of gold, to the process that leads, say, from having the valence shell electrons delocalised to the conduction of heat and electricity, there corresponds a (candidate) natural kind, gold, which consists of individuals/samples, that possess, *inter alia,* such delocalised electrons in their metallic structure.

It is the latter type of process however, the one along which organisms acquire/lose the determining (pathological) properties themselves, which poses problems from the point of view of criterion VIII to the GD kind. In these types of medical transitions, the corresponding kinds of objects (organisms) intuitively undergoing them seem to be characterised by other properties, more fundamental than the pathological 'determining' ones.[266]

What is to be said about the seemingly inescapable ontological gap between the two kinds? I shall first reckon with a host of preliminary reasons advanced in the literature in order to draw the line between phase and non-phase kinds. These reasons, I shall argue in §3.6.1, cannot show that gold satisfies criterion VIII in contrast to the GD kind, because they are either metaphysical intuitions (and intuitions can always be disputed) or they can be adapted to make the GD kind a non-phase one. Then I shall come back in §3.6.2 to

[265] *Cf.* Ellis (2001). At this point, Ellis introduces natural kinds of processes that correspond to the natural kinds of objects, in the sense that the former spring from the (manifestations of) the causal powers that the latter are associated with.

[266] The *bearers* of these transitory events seem to be members of a more profoundly grounded kind, say, human organisms. We need not dwell upon what exactly this more profound kind of organism should be. What matters is that the pattern of determining properties of the GD kind appears like a phase in a transition undergone by organisms that present a more stable/fundamental pattern of determining properties.

the discussion of the two senses of carving nature at its joints and argue that, unless the identity of kind members is taken into consideration, any further reasons adduced in order to situate the two kinds on different ontological levels cannot fulfil their purpose. Just like in chapter 2, I will discuss Jonathan Lowe's views, from a different perspective though, more suitable for the applied purpose of the present chapter.

§3.5.1 *TRANSMUTATION AND THE PHASE/NON-PHASE DISTINCTION*

The starting point of my analysis is represented by certain types of processes that gold can take part in. Exceptionally, gold is involved in changes in which its samples acquire and/or lose their determining properties. Such a process is the transmutation of gold into lead, an event during which the gold atom acquires three protons. One of the most important determining properties of the kind gold is thus modified. Accordingly, the kind members undergoing this transformation actually change their membership from gold to lead.

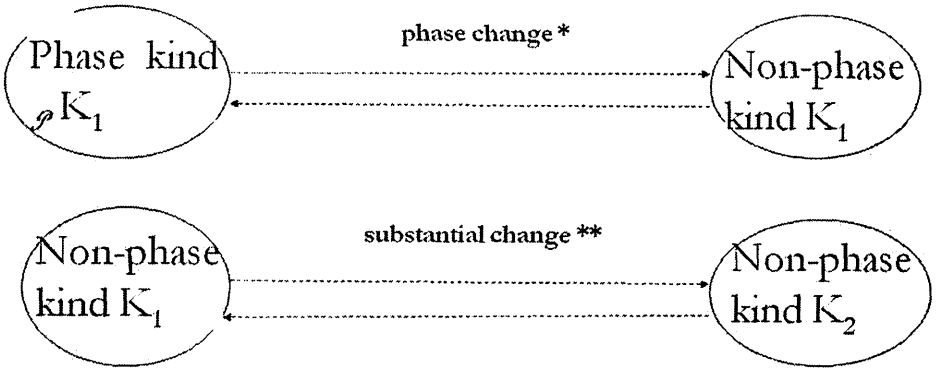
That gold samples can suffer such processes does not represent, as such, a reason to qualify it as a phase kind. What is crucial however is to evaluate whether their *non-phase* kind membership (if any) remains the same, is changed or *comes to* be possessed. Let me recall a few conceptual distinctions I have introduced in chapter 2.

Members of non-phase kinds could be involved in principle in processes in which their determining properties are changed. It is just that these processes are supposed to be *substantial* ones, that is, processes along which the initial membership in the respective non-phase kind is lost. They should be contrasted with the phase processes, that is, the processes marking out the transitions undergone by a non-phase kind in which individuals are not supposed to change their non-phase kind membership. More precisely, individuals should simply come to transitorily instantiate a certain pattern of properties that, as a *façon de parler*, might be called a phase kind.[267]

[267] For a discussion of the distinction between phase and substantial changes see Lowe (2001:182-184) and also Wilson (1999: 16-20)

So for instance the changes undergone by a caterpillar turning into a butterfly are supposed to be phase changes during which the (non-phase) kind membership of an organism is not changed. At most, the pattern of properties characteristic of caterpillars (which differentiates them from butterflies) might be said to delineate a phase kind. The dying of a butterfly, to the contrary, should qualify as a substantial change in which the membership in the initial non-phase kind is swapped for membership into a different non-phase kind.

***FIGURE* 15 PHASE AND SUBSTANTIAL CHANGES (the arrows indicate the direction of the transitions undertaken by kind members - from phase kinds to non-phase kinds and vice-versa in the case of phase changes, or from one non-phase kind to another in the case of substantial changes)**



* An individual *a*, member of the non-phase kind K_1 which underwent a *phase* change would not swap its non-phase kind membership in K_1 but would transitorily instantiate the phase kind φK_1

** An individual *a* member of the non-phase kind K_1 which underwent a *substantial* change would swap its non-phase kind membership from K_1 to another non-phase kind K_2

In fact, in such changes involving two different regna (biological and material) a change in the numerical identity of the individuals bearing the processes at hand is arguably also involved. This detail though will become more relevant later. What we need to hold on to at this stage is that only if the gold-lead transmutation was a substantial change would gold straightforwardly be a non-phase kind.

Now, we could usefully contrast the process of gold-lead transmutation with the sort of

process by which human organisms acquire/lose the determining properties of Graves' Disease. We could take for instance the scenario by which the organisms already possessing those pathological determining features lose them, say, under treatment. During such a process (which we could call 'the Graves process'), one of the most important determining properties of the GD kind is modified such that the kind members undergoing it actually change their kind membership (from GD to, say, 'healthy' organisms).[268]

On the one hand, we have, then, some human organisms in which the level of thyroxin in blood (initially dramatically high) decreases, due to a treatment with propylthiouracil. This drug inhibits an enzyme called thyroperoxidase, which has a crucial role in the synthesis by thyroid of its hormone(s). Once thyroperoxidase stops utilising H_2O_2 to produce I^+ in the apical membrane surface of the thyroid follicle, the residues of tyrosine, an amino-acid present in the composition of a dimeric protein (thyroglobulin), cannot be used for the assembling of thyroxine. On the other hand, we have some samples of gold in which the number of protons in nuclei is modified due to α-bombardment in a particle accelerator or by leaving them for a longer period of time in a nuclear reactor. By neutron capture and β decay at a small neutron flux, the gold atoms change into Hg, Tl and then Pb atoms.

Now, why claim that in the case of the Graves process we have a *phase* process - that is, a process due to which individuals move from a *phase* kind to another, (non-phase) kind[269] - whereas in the case of the gold-lead transmutation process, the transition is supposed to entail a change that is more profound than a *phase* one? In other words, why not say that the gold kind picks out a *phase* in the process leading to the lead kind? We will look first at four preliminary means supposed to draw the phase/non-phase distinction, which

---

268 In the minimal sense of organisms whose survival is not threatened, by Graves' Disease at least. We need not be preoccupied here with what the precise details of the kind instantiated after/before the organisms lose/acquire the determining properties of the GD kind represents (or with the kind that is presumably instantiated all along the sickness time-period). It is sufficient to say that it looks like a more fundamental kind, in relation to which the GD kind appears as representing a phase.

269 In the quasi-majority of cases, the non-phase kind that appears saliently at the end of the phase had been instantiated all along the phase change – see for instance the transformation of tadpoles into frogs. I do not wish to exclude the case in which, after a phase, the non-phase kind indeed *comes to* be instantiated.

appear to indicate that gold is a non-phase kind, in contrast to the GD kind.[270]

It might be said first that gold is much more stable than the GD kind and that stability is a mark of non-phase kinds distinguishing them from the transitory nature of phase kinds.[271] A second (related) reason that might be invoked is that the Graves process is quite frequent and easily replicable whereas gold-lead transmutation is artificial and not reversible.[272] A third reason might be that gold, in contrast to the GD kind, is an 'eternal' kind.[273] The chemical elements, along with kinds of particles in fundamental physics, one could claim, constitute the 'blocks' from which the universe is built, whereas diseases represent the paradigm of transitory, ephemeral phenomena. A fourth reason that could be adduced is that the Graves process is subject to natural laws of development. Just as there is a natural law (of development) that dictates the transformation of tadpoles into frogs - in other words, that establishes that tadpoles are a phase undertaken by frogs - the Graves process is governed by the laws of development (and regress) describing the interactions between human organisms and their environment. No such 'developmental' laws are in place for the gold-lead transmutation.[274]

Let us take all these reasons in turn. As regards the first one, it could simply be argued that stability is an epiphenomenon that does not even scratch the surface of the distinction between phases and non-phases. Processes could happen faster or slower. We might have patterns of properties that are instantiated for a very short time and yet seem

---

[270] These means are to be found in among others, in Lowe (2001), Millikan (1999), Boyd (1999) Robinson (2004). They only represent an illustrative selection from the variegated reasons advanced in order to draw the phase/non-phase distinction, in the discussions concerning phase kinds as such, sortals or the concept of substance. Other examples of features advanced in order to shed light on the vein substances/non-phase kinds/substantial sortals are: being subjects of predication and bearers of properties, being the subjects of change, possessing unity in our spatio-temporal framework, being ontologically basic, being the crucial entities in a given system, etc.

[271] See Lowe (2001:176) who claims that the phase changes should occur as effects of natural changes in the environment, and not due to drastic measures undertaken in laboratories. See also Robinson (2004) who approaches the phase/non-phase distinction as part of the broader discussion over the concept of substance and who mentions stability as well.

[272] *Cf.* Lowe (2001: 174-189)

[273] See Millikan (1999), Boyd (1999). The two authors discuss in their articles the differences between exact science kinds and special science kinds.

[274] 'A change to an individual substance S, of kind K, is a phase change for S just in case it is a change which things of kind K survive as a consequence of the natural laws of development for Ks' Lowe (2001: 186). Note that Lowe has different phase-conditions for objects belonging to different *categories* (the material, artefactual and biological). I employ here a few of his conditions for illustrative purposes. The laws of development are mentioned by Wiggins as well, as means to trace out phases (and phase sortals); see Wiggins (2001:143).

credible candidates for characterising non-phase kinds, as the case is for certain transitional chemical elements. Indeed, a proper example is unobtanium (Un^{346}) whose atoms decay in 4.8 microseconds. Conversely, that processes take place very slowly is no guarantee that the change they underlie is a non-phase one.

With respect to the second reason, it could be pointed out that the gold-lead transmutation can actually be reversed in laboratory conditions and it frequently takes place (in either direction, from lead to gold and from gold to lead) in supernovas. It is, after all, a natural process and the laboratory conditions in which it can be reproduced should not make us consider it artificial in any way.

The response to the third motive would be merely to say that gold is not an 'eternal' kind. The gold element was created about 100 million years after the Big-Bang in the stellar nucleosynthesis, like all the elements whose atomic number is greater than that of iron.[275] Indeed, it took about 100 million years for gold to be created, whereas for human organisms to emerge (and be capable of bearing the traits of Graves' Disease) some other 13,7 billion years were necessary. But no matter the time span that separates the emergence of the two kinds, what we have here is a difference of degree, and not an ontological difference.[276]

Finally, as regards the forth reason, a reply would be to emphasise that both the transmutation and the Graves process are subject to laws. What happens in the gold nucleus when it is bombarded with α particles, as well as what happens in an organism when the enzymatic pathways that assemble the chemical components of thyroxine are disrupted by propylthiouracil, can be subject to counterfactual descriptions and does not depend on the time and place in which the processes occur. That the laws describing the Graves process can be qualified as being laws of *'development'* whereas the laws of transmutation apparently cannot, has only an intuitive (and hence very weak)

275 See for instance Wallerstein *et. al.* (1997).
276 Mitchell (2000) makes this point when comparing biological with physical *laws*, but the 'created' nature of gold is just as relevant for the *kinds* discussion.

importance.[277]

The set of motives I have just invoked as potentially able to put gold and the GD kind on different ontological levels is just a representative selection of the variegated reasons advanced in the literature in order to draw the phase/non-phase distinction. All these reasons, however, could be similarly diagnosed - they are either metaphysical intuitions (and intuitions can always be disputed) or they can be adapted to the Graves process. In other words, they cannot show an ontological discrepancy between the two kinds. We seem thus to have a choice - either the GD kind is acknowledged on these grounds as a non-phase kind or we have to admit that gold is a phase kind.

§ 3.5.2 IDENTITY AND PHASES

I will now come back to the caveat regarding the carving nature at its joints. The crucial justification for the Graves process being a *phase* process, unlike the transmutation of gold into lead, I shall argue, could only be that the identities of the organisms that lose the determining properties of the GD kind are not influenced by this process, whereas the identities of the individuals members of the kind gold are indeed affected by the transmutation. As in chapter 2, I will discuss Lowe's stance on the issue of the phase/non-phase distinction, from a different perspective though.

Lowe describes in his (1998) four characteristics that processes undergone by kinds of 'stuff' should possess in order to qualify as *phase* ones. The example is the transition from water to ice.

'(1) The change is readily replicable, as a natural effect of an alteration in certain environmental conditions (2) the change is relatively gradual and continuous....(3) the change is naturally reversible...(4) during the change, no matter is lost or gained by the substance involved' (Lowe, 1998: 176)

Features (1)-(3) do not represent problems for our comparison. Features (1) and (3)

[277] And Lowe (the main proponent of the motive that the phase transformations are subject to laws) could not object that no laws are at work for the processes underwent by the GD kind members, since he rejects reductionism to the fizico-chemical level and admits himself that we could have laws with exceptions. See Lowe (1989: 153, 154).

actually figured in the set of preliminary reasons I have just discussed. Feature (2) represents a classical example of a condition framing, in an ambiguous manner, a metaphysical intuition. What does it mean for a change to be *relatively* or quite gradual and continuous, one could ask? No clear response could be provided, of course, and what makes matters worse is that Lowe acknowledges parenthetically that the ice-water change is discontinuous on a micro-structural level.[278]

The fourth feature is much more interesting however, because, in Lowe's metaphysical framework, it represents a criterion of identity. The 'preservation of matter' is for Lowe a condition that needs to be fulfilled by individuals from the category of 'stuff' (non-biological, non-artefact matter) in order to remain numerically the same across changes and processes.[279]

For Lowe then, numerical identity preservation is a necessary, yet not sufficient condition, for a process undertaken by an individual substance to be a *phase* one.[280] Conversely, for a process undertaken by an individual member of a certain natural kind to be a substantial one, the change of identity (i.e. its becoming a *different* individual) represents a sufficient, yet not necessary condition.

Now, if one denied, as I do, the relevance of conditions (1)-(3) for the phase/non-phase distinction in the context of our comparison,[281] it would follow that for a change to be a substantial one, the alteration of identity was a necessary and sufficient condition. Let us look at what Lowe has to say about the transmutation between lead and gold.

'Finally, consider the (hypothetical example of alchemic transmutation...if [alchemists] had succeeded, how should we have regarded the relation between a base metal, such as lead, and gold? Should we have

---

[278] Lowe (2001: 176)

[279] Lowe delineates three categories of 'objects', each of them associated with different conditions of identity (Lowe 1989: 8-15). As I mentioned in chapter 2, these conditions are such that they preclude any individual to trespass the limits of a category and yet remain identical with itself (see the example of any dying organism or the statue/material discussion, approached in Lowe, 1995). *Within* a category, insofar as the corresponding conditions are fulfilled, individuals can undergo all sorts of changes *including changes of their kind membership*, while remaining identical with themselves. In other words, Lowe adheres to the non-identity sense of carving nature at its joints. This aspect will become relevant a bit later.

[280] Lowe (2001: 174, 182)

[281] I do not exclude that these conditions might be useful to draw the phase/non-phase distinction in general, or in other particular cases. What I am precisely interested in however, as regards these conditions and any other, is whether they can reveal an ontological discrepancy between the two kinds under comparison here, from the point of view of criterion VIII.

regarded them as really being the *same* metal, differing only in *phase*, as water differs from ice? I suggest we should have done so only if the process of transmutation had possessed features (1) to (4). Suppose however that the process had possessed all of these features, but that in addition, scientists had discovered (as they now have) that gold and lead are composed of atoms differing in their nuclear composition. Would that additional discovery have warranted a reversal of the judgment that gold and lead are the same metal, differing only in phase? I think not. Thus I conclude that, as things actually are, what warrants the judgment that gold and lead are indeed different metals cannot just be the fact that there is no natural means of transmuting gold into lead which possesses features (1) to (4). It is true that the fact that lead and gold differ in their atomic constitution - though only in conjunction with other facts about atoms – explains *why*, as a matter of natural law, no such natural process of transmutation can occur. But it still seems clear that this fact about their atomic constitutions is not by itself the sort of consideration which should persuade us to regard gold and lead as stuffs of different kinds.' (Lowe, 2001: 177-8, italics original)

By Lowe's own standards then, the respective transmutation could qualify as a phase change if conditions (1)-(4) were met. Of course, Lowe does not actually hold that the gold-lead (or lead-gold) transmutation is a phase process. That is because, even if Lowe does not claim that any change in the identity of the samples involved takes place,[282] still, in his view, this process is not readily replicable as a natural effect of an alteration in certain environmental conditions, is not relatively gradual and continuous and is not naturally reversible.

If one denied, however, the relevance of features (1)-(3) for the phase/non-phase distinction, it would emerge that if and only if the identity of the involved samples was affected, the process of transmutation would qualify as a substantial, non-phase one. To put it differently, only if natural kinds were said to carve nature at its joints in the identity sense, would the gold kind not qualify as a phase kind and would hence be ontologically different from the GD kind (on the membership of which the identity of organisms could not possibly depend).

[282] The claim is not made directly. Lowe just insists that individuals' swapping of natural kinds, within a category, could take place without any change in their identity; see his discussion of substantial kind change and individual substantial change in Lowe (1998: 174,175,186). Since Lowe does not count the transformation of water into ice as attended by the losing of matter, presumably the gold-lead transmutation does not presuppose any loss of matter as well in his view. Note parenthetically that the identity at stake in the gold-lead transmutation is not numerical identity (since we are talking about samples of kinds of stuff), but diachronic identity in the broad sense, which is involved in our identifying of the *same* sample of a kind of stuff over time (see Lowe (1998:187)). Of course, my disclaimer concerning the absolute/relative identity distinction remains valid in this section as well – unless otherwise stated, any talk of identity (numerical or diachronic, more broadly) is about absolute and not relative identity.

If the *non*-identity related sense of carving nature at its joints was adopted, no other metaphysical elements or characteristics of the transmutation process could be invoked, I claim, without finding a 'correspondent' in the case of the Graves process. For instance, if what was invoked was that, in the case of the gold-lead transmutation, gold members lose some of their 'essential' properties,[283] then, if the 'essential' properties are the ones construed in the tradition of the Lockean primary qualities, the same could be affirmed about the Graves process. In other words, if such 'essential' properties were used as a criterion for drawing the phase/non-phase distinction, it would follow that either both processes were phase processes or both were non-phase ones.

At this point I find myself in complete agreement with Lowe who remarks in the above cited passage that '...lead and gold differ in their atomic constitution... this fact about their atomic constitutions is not by itself the sort of consideration which should persuade us to regard gold and lead as stuffs of different kinds.' Lowe is even more adamant in other places:

'Neither macroscopically observable phenomena, nor scientific information concerning the 'internal constitutions' of things, can resolve such issues for us in the absence of metaphysical guidance. Thus, even the fact, for instance, that the butterfly's DNA is the same as the caterpillar's cannot of itself show that the transformation of the latter into the former is not a change of substantial kind, any more than the fact that the atoms constituting lead and gold have different nuclear compositions of itself shows that the hoped-for alchemical transformation of lead into gold would not be a phase change'. (Lowe, 1998: 179)

On a different note, if it was invoked that in the case of the gold transmutation, gold members come to instantiate a different substantial universal,[284] then there would be no reason why the same could not be said about the Graves process. The latter process would indeed involve a more frequent change in the instantiation of different substantive universals than is involved in the case of the gold-lead transmutation. However, I would suggest, this 'frequency' aspect could not count as a motive not to accept such universals for the two kinds standing in the beginning and end, respectively, of the Graves process. The rationale I want appeal to is the merely the following: if Ockam's razor does not function when substantive universals for the kinds gold and lead are postulated, there are

[283] Of the sort Wilkerson (1995) for instance claims natural kinds have. See also §2.1 where such 'essential' properties are discussed.
[284] Lowe (1998)

no reasons why it should function in the case of biological kinds such as GD. Let me explain why I think this might well be the case.

I have mentioned in §2.2.1 *et passim* that it is a very contentious issue whether such substantive universals should be accepted at all. All the roles they are supposed to play – like being the ground for co-instantiation inductions or being truth-makers for natural laws – appear as redundantly filled in, because they can be fulfilled within a non-inflated ontology based on properties only.[285]Now, if in spite of Ockham's razor, substantive universals are still accepted in one's ontology of *natural* kinds, what restricts which kind members should instantiate substantive universals and which not? At any rate, it could not be the N&S requirement, as I have shown when discussing Ellis's work. Again, Lowe's position – the author who most prominently adheres to the view that natural kinds are based on substantive universals - needs to be discussed.

We have seen that, unlike Ellis, Lowe does not accept the N&S requirement. Lowe thus sees the existence of substantive universal-based kinds as compatible with the *members* of those kinds sharing a fuzzy set of properties. Therefore, on his own grounds, a kind such as the GD kind, which has cluster-like conditions of membership, could in principle be associated to a substantive universal. Moreover, if we take into account all the roles that substantive universals are supposed to play in the metaphysics of kinds, we can see that they can be successfully fulfilled by the putative substantive universal of the GD kind I hypothesize here. This GD substantive universal would be the truth-maker of the laws in which organisms with the characteristics of Graves' Disease are involved. Such an universal would also explain why certain determining properties are co-instantiated by the members of the GD kind. It would also support the counterfactual scenarios in which the organisms at hand could be meaningfully invoked and ultimately, for the exceptions on the level of particulars that are commonly dealt with using *ceteris paribus* clauses attached to first-order quantified statements (in the literature on special science, in general, and on medicine as well).[286]

More plainly put, my suggestion is as follows. If the only criteria for a group of individuals

[285] See Heil (2006), Bird (2007)
[286] Lowe's exposition of the advantages of embracing substantive universals can be found for instance in Lowe (2006: 26, 132, 133, 147)

to instantiate a substantive universal are that they co-instantiate a certain set of properties and are commonly engaged in nomic behaviour (*ceteris paribus* or otherwise), and also if the N&S and the non-overlapping requirement are rejected or shown to be unjustified,[287] then there is no reason why substantive universals could not actually be associated with a huge number of kinds, the GD kind included. Once Ockham's razor is stopped from functioning at the initial stage when the very existence of substantive universals is introduced (vis-à-vis the kinds in the exact sciences), it could not be brought to function and limit such a 'multiplication' of substantive universals (in the special sciences domain, medicine included).

What could stop this multiplication of substantive universals? To be sure, a means to stop it (the only one I can see in fact) is linking the instantiation of a substantive universal to the identity of the respective kind members. [288]

Think for instance about the many different classifications and kinds that any single human organism could be said to instantiate. Graves' Disease is one of them, but we could also mention here, say, being a child, adolescent, elderly person, or, from the range of diseases, suffering from Influenza pneumonia, peptic ulcer or glomerulonephritis (to cite some very common pathological conditions). All these classifications and the attending kinds of organisms smack of phases and *phase* kinds, as it were. Their possible association with substantive universals could be dismissed (and accordingly the phase/non-phase distinction could be employed to rule them out as *natural* kinds, according to criterion VIII) if we conceded that the identity of any particular organism member of a natural kind depends on its instantiation of the respective substantive universal.

Of course, what I am referring to at this point is the non-overlapping requirement, *in* the context of the identity sense of carving nature at its joints. All the above-mentioned kinds overlap each other – not all adolescents suffer from glomerulonephritis (even though some of them do) and not all elderly people suffer from peptic ulcer (even if some of them

[287] As I have argued in chapter 2

[288] Indeed, this was the construal of substantive universals in Aristotle and his medieval followers; see Gobbo (2000: 84), Dahl (1997: 233), Ayers (1991: 20). Oderberg is the only modern author who holds that kinds are substantive universals and claims explicitly that the phase/non-phase distinction should be drawn using the identity of kind members; see Oderberg (2007: 268), Oderberg (1996: 150-152).

do). If the identity of any particular was made dependent on its instantiation of a substantive universal, at most one of these kinds would have to be accepted and all the other putative substantive universals could be dismissed, on grounds of the *tertium non datur* law of identity.[289]

Lowe would not accept this solution though. We have already seen that in his construal of kinds, an individual could swap natural kinds even if its identity remained unaltered. In other words, an individual could come to instantiate different substantive universals while remaining numerically the same. For Lowe then, natural kind membership should not have an identity bearing on kind members. Nevertheless, if this is the sense in which kinds carve nature at its joints, then the ontological difference between gold and the GD kind arguably disappears, in the context of the substantive universals discussion and that of criterion VIII as well.[290]

§3.6 AGAIN, ON CARVING NATURE AT ITS JOINTS (AND ON CRITERION IV)

One important aspect of my argumentation was set out in the discussion of the carving nature at its joints intuition(s) in §1.5. The guiding line I have adopted vis-à-vis the intuitive burden surrounding the kinds theme was that we have two choices when adverting that natural kinds 'show what things are' or delineate 'the fundamental divisions of nature', etc. Either kinds are considered natural when the kind membership has a bearing on the identity of its members, or kinds are considered natural, by employing all sorts of other reasons, irrespective of their bearing on the identity of their members.

The point I have tried to make in chapter 2 and the present chapter was that if the non-identity related sense of carving nature at its joints was adopted, then all arguments and criteria for gold being a natural kind can also be applied for the case of the GD kind.

---

[289] See my discussion in §2.1.2 of the 'paradoxes' of criss-crossing.

[290] Authors who believe that identity is crucial for drawing the phase/non-phase distinction, are Wilson (1999: 20) Hirsch(1982: 52-57) Wiggins (1980), Oderberg (2007). Not all of them see kinds as substantive universals. Some epistemic problems attending their position will be looked at in the next section.

On the other hand, if the identity related sense of carving nature was chosen, then an ontological difference between the GD kind and gold kind could arguably appear in the following sense. Given that the identity of the organisms could not possibly depend upon their diseased traits,[291] if the identity of gold samples depended upon the microstructural characteristics of their kind, then the two kinds under comparison would indeed be situated on different ontological levels. This represents an extended formulation of the identity-proviso that I announced in Introduction to the present thesis. In this section, I simply want to point to a host of reasons due to which such a construal of kind membership is very difficult to justify, from an epistemological point of view.

It should be said that most kind theorists do not discuss the issue of the identity of kind members or prefer to adhere explicitly to (what I have called) the non-identity sense of carving nature at its joints. Nonetheless, some authors do claim that the kind membership does play a role in the identity of kind members.[292] Their position faces some serious epistemic difficulties, as I have said, and I will explain in the following why.

We should recall first of all that, aside from the entire issue of intuitions regarding the category of natural kinds, to consider that kinds carve nature at its joints in the identity sense is to bring up criterion IV as a requirement that should be satisfied by any candidate natural kind. Criterion IV (the identity requirement) states that for any individual *a* - member of the (candidate) natural kind K characterised by the determining properties [p_1, p_2, ...p_n] - were *a* to stop instantiating them, *a* would turn into a different individual *b*.

Then, we should clearly distinguish between two aspects that could be addressed at this point. On the one hand, one could ask whether it is intelligible to generally consider that such identity-influencing properties exist. On the other, one could look into why these properties should be the *determining* properties that characterise the commonly accepted kinds in the exact sciences. The former enquiry concerns a metaphysical issue that I will not much refer to in this section. For my part, the existence of such properties should be

[291] I take this affirmation as uncontroversial.

[292] Some of the proponents of criterion IV are Wiggins (2001), Brody (1974), Denkel (1996), Elder (2007). The present section will synthesize a classic critique, to be found in an extended version in Mackie (1994) and Sidelle (1992) for instance.

acknowledged. At any rate, I will simply accept here the metaphysical thesis that the identity of individuals depends upon some the properties they instantiate (i.e. the thesis that there are no bare individuals).[293] It is the latter question, posing an *epistemic* challenge, which is the subject of this section.

Now, could we justifiably say that the *kind* properties play such an identity role – where the kinds in question are the ones commonly accepted in the exact sciences, in virtue of certain patterns of properties? That is, why should the identity of *a* depend upon the properties [p1, p2,...pn] characterising its kind K? These are at base determining properties and we have seen that the commonly accepted kinds in the exact sciences - for which K stands here - satisfy in addition various other criteria. Indeed, it is precisely these determining properties – constituting necessary and sufficient conditions (for delineating the division of nature corresponding to K), characterising K as intuitively a non-phase kind, putatively precluding individuals like *a* to simultaneously instantiate any other kind K* that was not in a species/genus relationship with K – that should be proven to have an identity influence over *a*. As I have tried to show in the previous section, in order for the N&S requirement to be justifiably adopted, *all* of the $[p_1, p_2, ...p_n]$ properties (circumscribing the non-fuzzy division of nature corresponding to K) should have an identity-influence over *a*. The same could be said with respect to the non-overlapping requirement and the non-phase requirement, even if, as I have remarked, the latter two criteria are not logically connected with the former.[294]

Yet, there does seem to be a distance between paying full heed to the sort of properties that science tells us about and drawing conclusions about the identity of kind members, even if the kinds in question belong to the exact sciences. Paul Teller, from a rather neutral position, nicely expresses this difficulty.

....I call this the problem of co-variance or co-occurrence of properties: given just the assumption that two

---

[293]For discussion, see Chisholm (1967) Mackie (2006)

[294]That is, we could admit that (candidate) natural kinds have a cluster of characterising determining properties and yet demand that no individual instantiated two clusters belonging to two different kinds which were not hierarchically positioned, or that the non-phase membership was lost if a certain proportion of a (candidate) natural kind cluster was lost by an individual, respectively. If this weaker, cluster alternative was adopted, a certain proportion of the (perfect) clusters of the divisions of nature in exact sciences would need to turn up not being identity-connected, in order for the hierarchy and non-phase requirements to be ungrounded.

properties co-vary in their extension over the range of all possible worlds, it is quite consistent to assume that a thing has both of these properties in one possible world while that very same, numerically identical thing, has neither property in another possible world...the generalisations of contemporary science....do no more than specify co-occurrence of properties...atomic physics describes something called the 'transmutation' of carbon into nitrogen....*This generalisation can be glossed by saying that one existent, the carbon atom, ceases to exist and something else, a nitrogen atom, comes into existence; or by saying that a carbon atom changes into one atom of hydrogen while retaining its numerical identity.* (Teller, 1975: 238, 239, italics added)

It is exactly the difference between being a determining property and having a bearing over the identity of kind members that is underlined by Teller.[295] Modern science says nothing about the identity of individuals and hence is consistent with the commonly accepted kinds not satisfying criterion IV (to use my terminology).

We thus have a possible ontology, alluded to by Teller, which is worth looking at when discussing the epistemic problems of the identity of carving nature at its joints. Were we to draw it in a bit more detail, we would not have to squabble with the empirical information that science provides us regarding generalisations about co-variance or successions of properties. We would only have to accept that the identity of chemical elements and chemical samples, for instance, does not depend on the respective atomic numbers, that the identity of electrons does not depend upon their charge, that the identity of protons does not hinge on their having two down-quarks, and so on.

Recall that the metaphysical thesis according to which some properties possessed by individuals have an identity influence is not under dispute here.[296] Hence, the identity of the individuals involved in processes and transitions could admittedly depend over some of their determining properties, but not on the ones commonly taken to characterise and singularise the commonly accepted kind of the exact sciences.[297] For instance, the identity of gold samples could depend on the constituting atoms possessing, say, at least 73

---

[295] Even Ellis – who, as we have seen, should be directly interested in construing the respective properties into an identity key - expresses the same perplexity as Teller's. See Ellis (1999:67,68) (2001:239) (2002: 12, 2-5, 238)

[296] My discussion goes thus on the line set out by Mackie, in her (1994: 311, 312, 332)

[297] I have mentioned that the identity of the samples involved in the changes studied by the exact sciences might not depend on their determining properties (their 'composition' to take up the classic idiom) but on some of their superficial properties. Lowe argues thus that it is at least conceivable to take the identity of water samples as depending over their perceptible characteristics (see Lowe 2007b).

protons in their nucleus, having at least 2 electrons in the 5^d sub-orbital and an ionisation energy of at least 1.5 in the Pauli scale. A gold sample would not thus change its identity unless it stopped instantiating these properties – which obviously do not match the determining properties characterising the commonly accepted kind gold.

One might of course wonder whether we could still maintain that the (candidate) natural kinds are the ones that have identity - influencing properties and modify our framework of kinds in the exact sciences. To stick with our example, instead of gold being considered a (candidate) natural kind for instance,[298] we could have as candidate the much broader kind with at least 73 protons in their nuclei, having at least 2 electrons in the 5^d sub-orbital and an ionisation energy of at least 1.5. Such a kind would have as members, besides 'gold' samples, all the chemical samples that are commonly classified as Tantalum, Tungsten, Rhenium and so on, going up from the elements with atomic number 73 towards gold.

The possibility of modifying our framework of kinds on this track (such that criterion IV was respected) does not seem to have great prospects though. Once the distance between being a determining property and having an identity-influence is clearly marked out, a myriad of other scenarios could be construed in order to build up such alternative kind frameworks. What could debar us from taking instead, as a (candidate) natural kind instantiated by the gold samples, the division of nature whose members have at least 46 protons in the nucleus and a relative atomic mass is at least 106.42, or the division with the first molar ionisation energy at least 745.5 $kJ \cdot mol^{-1}$ and an atomic radius of at least 128 pm?[299]

How broadly or narrowly should these kinds (which, evidently, do not look like the classical kinds accepted in present science/philosophy of science) be construed, in the pursuit of respecting criterion IV, appears accordingly to be largely a matter of choice, due to our epistemic ignorance. In a quite broad scenario, we could only have three (candidate) natural kinds – the material, the biological and the artefactual - if we accepted the vision of identity conditions that Lowe proposes for instance and we

[298] As it is commonly construed in nowadays science, i.e. being circumscribed by the determining properties described in § 3.1.2

[299] See Mackie (1994: 330, 331)

demanded in addition that (candidate) natural kinds had an identity influence. In the broadest possible scenario, we would only have one kind that the exact sciences are concerned with – the material kind, to be opposed, by those who wish, to the immaterial, intelligible one.

Note that this possibility of broader kinds, envisaged in order to keep up with criterion IV, would not entail any direct, metaphysical consequences over how we construe *individuals*, or more precisely, the separation between individuals. A world containing only one kind, or three kinds, as the one I have suggested above, would be consistent with having exactly the same differentiations between individuals as the ones commonly acknowledged. It is just that these individuals would remain numerically the same across a larger number of changes. For example, in the scenario with only three kinds I have envisaged,[300] any individual lead atom if transmuted into gold, would remain numerically the same. Similarly, any individual gold atom, if transmuted into lead, would not change its identity, if all the other conditions beside maintenance of their *material* status (i.e. spatio-temporal continuity, continuity of causal history, etc.) were also in place. Such atoms would become different individuals only in the (counterfactual) scenario in which they were not material any more but biological or artefactual.[301]

These scenarios with only one or three existing kinds are extreme ones and are brought into the discussion here for illustrative reasons only, following the argumentative line of our epistemic ignorance about identity, which Teller alludes to. However, we need only bear in mind that given the distance between being a determining property and having an identity influence (and also, depending on whether one accepts or not criterion IV) we have two possible scenarios. Either the kinds commonly accepted in exact sciences cannot be justifiably said to fulfil criterion IV (and hence be natural)[302] or alternatively, a myriad of other kinds fulfilling criterion IV could be envisaged (without us having the epistemic means of adjudicating between alternative kind frameworks).

Yet, as I have said, some kind theorists who stick to the identity sense of carving nature

---

[300] .... following Lowe's proposal over the identity conditions for particulars. Obviously, Lowe would not agree that the view that out there we only have three kinds, since he does not embrace criterion IV.

[301] *Contra* Lowe (2001:187) who claims that, in his terms, Spinozism about kinds entails Spinozism about individuals. For a more extended discussion of Lowe's views on this matter, see § 2.3.

[302] Therefore, we could only accept them as natural in the *non*-identity sense of carving nature

and, at least in their examples, cite as natural the kinds commonly accepted in the exact sciences.[303] They appear, that is to say, to consider the epistemic problem I have just described as solved or as not even a problem. Their argumentation proceeds over two distinct (and yet related) lines.

They argue, on the one hand, that if the characterising properties of kinds did not have an identity bearing, then, on an epistemic/conceptual level, we would not be able to discern among the *individuals* of our world. On the other hand, they hold that we could not otherwise maintain, on a metaphysical level, that the differences between individuals are in the way they are. That is to say, we would either have to acknowledge our epistemic and conceptual ignorance in distinguishing *individuals* and/or we would have to accept, as far as the metaphysical argumentation is concerned, a world in which the distinctions between individuals were blurred.[304]

I do not propose to discuss here in any detail their views, but only to indicate that the epistemic dilemma I have portrayed above cannot be very easily circumvented by such *modus tollens* - type argumentations. What I want to suggest, at a minimum, is that the purportedly unpalatable consequences of the two lines of argumentation do not straightforwardly follow from a denial of the commonly accepted kinds satisfying criterion IV.

That is because, arguably, we are able to distinguish among individuals (individuals *as* members of one kind or another included) without knowing *which* properties have an identity influence. Prior to the identification of individuals and kinds is the identification of properties.[305] Once properties are pinned down, the differentiations between individuals should become salient. Even in the case of perfectly similar individuals from a qualitative point of view, the spatio-temporal coordinates should stop us from (what we

---

[303] Wiggins for instance cites science and the laws of development discovered by scientific enquiry that should help us distinguish between kinds. *Cf.* Wiggins (1980: 77-83 *et passim)*; see also Weatherson (2002) for a discussion.

[304] See for instance Oderberg (2007: 117-125, *esp.* pp. 123, 125), Hirsch (1997: 7-15). Lowe employs the same type of argumentation, not in order to defend criterion IV, but to justify his metaphysical categories, which, as I have noted in the previous section, are different from his natural kinds; see Lowe (1989: 56, 57), (2009: 30)

[305] See Dupre (1993:80)

could call) inappropriately applying the converse of Leibniz's law.[306] Once properties are manifest, kinds are also recognisable and, *ipso facto*, their members are as well. A book is in the first instance identifiable via its properties. By considering different parts of these properties or even all of them, various kinds that a book can be taken to be a member of can then be grasped. Accordingly, the book in question can be viewed *as* a kind member and distinguished from other kind members.

On the other hand, it is arguably a *non-sequitur* to infer that, were the kinds commonly accepted in the exact sciences not to satisfy criterion IV, we would have to accept a modified reality of individuals. Take the scenario I have just discussed, in which we would only have three (natural) kinds with identity bearing – the biological, the material and the artefactual. To repeat, this scenario seems perfectly consistent with having exactly the distinctions between individuals in our ontology. No empirical information that science provides us would need to be overlooked if we wanted all the chemical elements to be members of the *material* natural kind. We would have, in Teller's idiom, the same generalisations about co-varying properties.[307] Why should we have the distinctions between any two atoms blurred if these atoms would always remain identical with themselves as long as they remained material (and are spatio-temporally continuous, etc.) and the material kind would be the only available one for them, as satisfying criterion IV?

It should be said that this entire discussion is framed in the literature in a rather idiosyncratic way, in that various other notions are introduced. Examples here would be sortals (where sortals could be concepts, terms or universals), principles of individuation (which could be epistemic or metaphysical), relative and absolute identity, criteria of sortal, synchronic and diachronic identity, trans-world identification, conceptualist vs. realist vs. conceptual-realist frameworks, etc. However, on scrutiny, the core of the discussion is close to what I have just framed above, in admittedly simple terms. Let me give just one example – Wiggins's discussion of sortals in his (1967), (1980) and (2001).

---

[306] See Black's famous example against the converse of Leibniz's law – the identity of indiscernables, discussed for instance in Noonan (2006). See also Grady (2007, section 2.1) who distinguishes, in the context of the so-called 'criteria of identity' construed in the cognitive/epistemic sense, between necessary and sufficient, and sufficient only criteria. The latter could be based on spatio-temporal coordinates only and could suffice for us to identify individuals.

[307] Or the same particulars, were we to speak about samples of 'stuff' (non-countable kinds) and not individuals *per se*.

Wiggins is a defender of what I have called criterion IV. His main line of argumentation proceeds as a reaction to Geach's thesis on the sortal relativity of identity statements. According to Geach, statements such as '*a* is the *same K* as *b*' do not entail that *a* and *b* are the same individual in the sense suggested by absolute identity, i.e. do not entail that '*a* is a K and *b* is a K and *a=b*'.[308] Geach's sortal relativity thesis attends his more general claim that the notion of absolute identity is not intelligible – roughly, that any a=b identification is relative to the resources of a certain language L, such that in a different language L* a=b does not hold any more (in the sense that a and b are not any more indiscernable).[309]

Understandably, Wiggins wants to reject both the thesis of relative identity and the one on the sortal relativity of identity statements. If both the former and the latter were correct, kinds could not possibly be said to satisfy criterion IV. Interestingly, the latter does not entail the former, in that one could hold the sortal relativity and accept the existence of absolute identity.[310] What sortal relativity would say in this case is that two numerically distinct individuals might represent the same kind member.[311] Accordingly, to reject the sortal relativity, while admitting that absolute identity exists, is simply to say that the same kind member must be one and the same individual. And, it does appear, from the assumption that the same kind members are the same individuals, there does not follow that the identity of individuals depends on the kind they instantiate.[312]

Now, Wiggins does acknowledge the separation between the two theses. His way of acknowledging it however is to indicate that there is a difference between denying that "a is the same K as b" entails "a=b" and denying that "a=b" entails that "a and b are of the

---

[308] That is, Geach denies that for any sortal term K, the expression '***a*** is the ***same K*** as ***b***' will be satisfied by a pair <***a***,***b***> only if the I-predicate of ***L*** is satisfied by <***a***,***b***>; see Noonan (1996) and also Geach (1962:152) *apud* Shoemaker (1970:532)

[309] More precisely, the claim is that we cannot make sense of an I-predicate standing for an equivalence relation (a relation that was symmetric, transitive and reflexive and respected Lebniz's law) without relativising it to a certain language.

[310] See Noonan (1996)

[311] See Geach's famous example of the cat Tibbles, whose tail undergoes strange spatio-temporal changes, which is analysed for instance in Noonan (2006)

[312] *Cf.* Noonan (1981: 264)

same kind K".[313] In other words, for Wiggins, to affirm that absolute identity exists is to say that something like criterion IV holds.

To be sure, one could point out that there are other ways of expressing absolute identity, via reflexivity and Leibniz's law *simpliciter* for instance, without bringing in sortals and kinds. And of course, as Wiggins's critics are keen to underlie, even if we come to terms with the understanding of absolute identity as involving sortals (which are the conceptual counterpart of kinds, in Wiggins) we have no warrant that the respective identity-bearing kinds can be known or that they correspond in any way to the commonly accepted kinds in science.[314] Another interesting aspect is that, once sortals are invoked in connection to the absolute identity of kind members, the separation between the two theses I have mentioned above becomes almost superfluous - the sortal relativity of identity statements can be shown to be trivially false in a few steps that involve a sortal variant of Leibniz's law.[315]

Wiggins has certain means of defense, which are complicated, interesting and provoking but do not alleviate the epistemic worries about sortals. Thus, Wiggins brings up certain so-called principles of individuation that are mainly related to our cognitive capacity to discriminate between individuals when employing sortal concepts.[316] The claim is that an individual falling under a certain principle of individuation could not escape its bearing without ceasing to exist. To the charge of anthropomorphism,[317] Wiggins responds by negating the distinction between conceptualist and realist ontological stances and situating himself in a middle line. As a conceptual-realist, he attributes to our cognitive activities the capacity to single out the identity-related divisions of nature, commenting in addition that these divisions are mind-independent and yet, they appear to us as they

---

[313] I indicate here Wiggins's distinction between what he calls the thesis R (affirming the sortal relativity) and the thesis D (affirming that any absolute identity is to be *elucidated* by the employment of a sortal). See Wiggins (1980: 48-50) Wiggins (1967: 1) and also Shoemaker (1970: 529,539), Noonan (1981: 260,261,264) for discussion

[314] See Mackie (1994: 330, 331)

[315] The steps are the following i) *a* is the same *K* as *b* (Assumption) ii) *a* and *b* are both *K** (Assumption) iii) if *a* and *b* are the same *K*, then they have all of the same properties (Leibniz' Law) iv) *b* is the same *K** as *b* (Reflexivity) v) *b* has the property of being the same *K** as *b* (from iv) vi) *a* has the property of being the same *K** as *b* (from iii and v) vii) *a* is the same *K** as *b* (from vi) See Grady (2007: section 3.2) and Wiggins (1980: 22) *apud* Noonan (1981: 262)

[316] See Wiggins (1980, chapter 4) *apud* Mackie (1994: 321, 322)

[317] See Lowe (2003: 818-820) and also Yablo (2003)

do because our concepts are in the way they are and not otherwise. Objects, says Wiggins, do not 'single out themselves' – they are the fish that do not slip through the mesh that we cast.[318] The matters are complicated by the fact that for Wiggins one principle of individuation can be shared by different sortals and it is not very clear which sortals share which principles.[319]

As I have said, it is not my purpose here to explore in more detail the niceties of identity and identity-influencing properties. In the context of our comparison, whether the determining properties of the gold kind, as commonly recognized, have an identity bearing over gold samples, is an issue that I will remain neutral on. Overall, what I just wanted to suggest was that whether criterion IV and the identity-sense of 'carving nature at its joints' could be employed in order to deem as natural the commonly accepted kinds in the exact sciences is a very complicated affair.

§3.7 FURTHER REMARKS ON MEDICAL KINDS

My purpose in the first three chapters of this thesis was to indicate, via a representative comparison, that as regards the life-threatening conditions in somatic pathology, there are no convincing reasons to affirm that medicine lacks a determinate structure of natural kinds, in the sense in which exact sciences can be said to possess such a structure.

The main direction of the entire enquiry was laid out in the realism/nominalism discussion of §1.4. I now wish to recapitulate my rationale, delineate the precise sphere of my conclusions and add some comments about the picture of natural kinds that can be inferred from my discussion.

---

[318]Wiggins (2001: 159) *apud* Lowe (2001:817)

[319]Mackie observes Wiggins's hesitations when it comes to the sortal 'animal' for instance and argues that the only characterisation of what a principle of individuation corresponds to on an ontological level (patterns of activity and development) is insufficient to make sense of which biological sortals share which principle of individuation; see Mackie (1994: 325, 326). Parenthetically, almost all of Wiggins's commentators recognize the originality of his work but remark his convoluted style of writing; see Lowe (2003: 816), Shoemaker (1970: 529) and Noonan (1981: 260)

While remaining agnostic about what natural kinds really are, one can safely assume that natural kinds (if nature has indeed such a metaphysical category) are to be found in the exact sciences. In this sense, gold can be construed as a classical natural kind. Using all criteria of membership advanced in the literature, one can compare then, from such an agnostic position, a medical kind and a classical natural kind. There are various reasons why such a comparison might appear as unintelligible and one criterion of membership is deeply involved in the intuitive obstacles that have to be dealt with. I have used criterion IV (the identity requirement) in order to frame a basic divide between the various substantiations of the intuitive assumption that natural kinds 'carve nature at its joints'. The respective divide represented a guiding line for my analysis. The identity of particular kind members, I have further argued, has to do not only with how one situates oneself against the claims that natural kinds 'show what things are', delineate 'the fundamental divisions of nature', etc. It is also involved in the argumentations surrounding the other criteria of membership (VIb, VII and VIII, that is, N&S, non-overlapping and non-phase requirements, respectively) and how the two kinds under comparison fare with regard to them. My conclusion was that unless the identity-related sense of carving nature at its joints is adopted – and serious epistemic difficulties stand in the way of this option – no ontological gap can be proven between the two kinds.

Some further clarifications need to be added to this conclusion. I should stress first that the use of criterion IV in such a pivotal role in my analysis did not represent a deviation from the general line of neutrality vis-à-vis the kind realism discussion that I have adopted - none of the two branches of 'carving nature at its joints' was given priority in any way. Another (related) qualification that should be emphasized is also concerned with the intuitive burden associated with natural kinds and especially with chemical elements. Irrespective of what arguments could possibly be adduced, one still feels that gold and all the chemical elements present features that situate them above the level of diseases. However, my conclusion was not that there is no difference at all between the (members of) GD kind and the gold kind. Throughout this chapter, I have noted several differences between the members of the two kinds in question. Let me recapitulate them.

The members of the gold kind exhibit behaviours that are described by strict laws whereas the members of the GD kind are arguably governed by ceteris paribus laws. Even

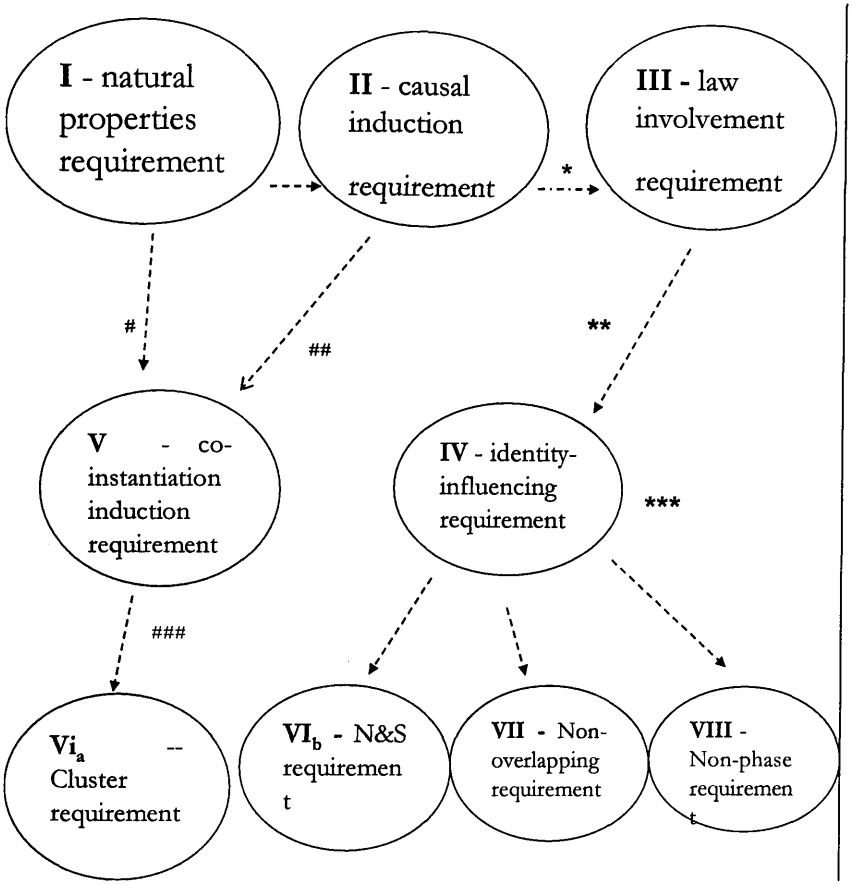
though both kinds are delineated by patterns of co-instantiated determining properties, in the case of the gold kind such properties explain almost all of the causal interactions in which its members are involved, whereas in the case of the GD kind the pathological properties explain only a few of the behaviours and superficial properties exhibited by the human organisms. Moreover, the determining properties of gold are much more 'stable'. That is, their unity, their co-instantiation in particulars is much harder to disrupt than the unity of the determining properties of the GD kind. Issues concerning supervenience and reductionism between the biological and the physical, which I could not discuss here, could also be brought in to underlie differences between the members of the two kinds.

Nonetheless, all these differences - insofar as one looks specifically at the metaphysical category of (natural) kinds and not at the (related) categories of causation, laws, properties, universals, etc. – represent differences of degree. They do not amount to an ontological gap when the criteria of (natural) kind membership are carefully reckoned with.

My discussion of the criteria of membership was not a simple application of the requirements involved. Along the analysis, I have discussed critically several criteria of membership (namely criterion VIb, VII and VIII). More precisely, I have shown in chapter 2 that if they are not read in conjunction with criterion IV (and if criterion IV is shown not to be satisfied), they are groundless. No matter what metaphysical elements are invoked to back up these criteria (substantive universals, essences, laws) if the identity of kind members was not brought in, then they could only be involved in circular reasoning that would arbitrarily characterise the category of natural kind.

After my discussion, the ontological dependency between the criteria of membership, which were preliminarily presented in chapter 1, can now be reformulated using the following diagram. In so far as my enquiry results in general conclusions about the character of natural kinds, the below representation of the dependencies between various features attributed to the metaphysical category of natural kinds would represent my view on how should any discussion about the 'divisions of nature' be pursued.

FIGURE 16 ONTOLOGICAL DEPENDENCIES BETWEEN CRITERIA OF NATURAL KIND MEMBERSHIP



* The sort of properties characterizing a kind which fulfills criterion III would have to be at least natural (i.e. to satisfy criterion I) and at least determining properties (i.e. to satisfy criterion II). In other words, criterion III entails criteria II and I.

** The most plausible candidates for identity influencing properties are the properties involved in laws.

*** Only if criterion IV was satisfied kinds could be attributed necessary and sufficient conditions of membership, would always be in a species-genus hierarchy if overlapping and would be differentiated from their phases. In other words, satisfaction of either criterion VIb, VII or VIII entails that criterion IV is also satisfied.

That kinds participate in co-instantiation inductions means that they are characterised by natural properties.

Co-instantiation inductions are dependent upon prior causal inductions.

The existence of clusters resulting from 'the causal structure' of the world represent the metaphysical ground for co-instantiation inductions. That is, satisfaction of criterion VIa represent a sufficient condition for satisfaction of criterion V.

My primary interest is however in medical kinds, against the backdrop of the (already very rich) literature on kinds in the exact sciences. I should thus finally state that, on the non-identity-related sense of 'carving nature at its joints', I adhere to (and hope to have argued intelligibly in favour of) a framework of natural kinds of diseases that suits, I believe, medical research and practice. In medicine we do have determining properties that are explanatory powerful. They are to be found in clusters that are instantiated by diseased organisms. The clusters can overlap each other and their limits might be fuzzy, in the minimal sense in which for some somatic (life-threatening) illnesses there are no necessary and sufficient conditions for diagnosis. In all the branches of somatic medicine, genetic classifications criss-cross with infectious, immune or morphopathologic schemes of arranging diseases and syndromes. Organisms can instantiate multiple diseases and the

treatment of the symptoms they exhibit is directed at eliminating the determining properties that produce them.

To stress, that certain criteria of natural kind membership are failed by disease kinds should not worry us in the least. These criteria would have ontological significance only if they were *justified* and I have shown that in their current form they are not. As to the very role played by classifications in medicine, it is worth adding that, for instance, the fact that the non-overlapping criterion is not considered in organizing disease entities is rather fortunate. As I said, in medicine, classifications criss-cross each other depending on what determining properties are under focus, be they genetical, immunological, anato-pathological or immunological. These classifications may coincide and gather under their heading the same groups of human organism with the same symptoms; but they do not always do so and this can only contribute to the richness of explanatory hypotheses in use for explaining diseases.[320]

One aspect I would have liked to discuss more was the issue of reductionism/supervenience of causal powers. Since my model has at its centre determining properties (as causally efficient properties) and these determining properties are clustered in individuals as a result of the causal structure of the world, causation is in a sense central to my approach. I have explained however in §3.1 and §3.2 why I believe that the niceties of the causation-discussions do not necessarily have to be brought in when discussing the theme of natural kinds. One can define the determining pathological properties rather loosely, as those biological, microstructural properties that can act as causes for observable pathological properties and behaviours (the symptoms). The existence of such properties can be inferred by a double negation, so to speak – it is

[320] For instance, immunology has quite a different classificatory system than antomo-pathology or infectious diseases. What happens in practice is that these classificatory systems, which obviously overlap each other, are used complementary in order to explain particular diseases, particular patients and their symptoms. Take the case of glomerulonephritis, for instance, which is categorised into several different pathological patterns vis-à-vis the level of inflammation in glomeruli, the immunoglobulin deposits present in the mesangial matrix or against the basal membrane, the presence of fibrin deposits (crescents) in the Bowman space, etc. (see Kasper *et al* 2008). Hence we have, as disease kinds, infectious glomerulonephritis (resulting from classifications on infectious causal grounds) membraneous glomerulonephritis (on immune grounds) crescentic glomerulonephritis (on morpho-pathological grounds), etc. These kinds criss-cross each other. Far from being a problem, the overlapping classifications corresponding to these kinds offer multiple explanatory hypotheses that concur in the diagnosis and treatment of any patient with fever, hematuria, proteinuria, nephronic or nephritic syndroms, etc.

impossible that processes in medicine happen by chance, accidentally. Hence, classifications that follow causally efficient bases shared by organisms are closer to being natural kinds than classifications based on other grounds (say, astrological influences, or belief in plotinian universals, as it was the case for syphilis or epilepsy). In fact Boyd, the main proponent of the model I have (roughly) adopted, does not discuss in any detail causation at any point of his argumentation about the homeostatic clusters to be found in the special sciences.

The issue of reductionism/supervenience for causal powers would have deserved more attention though, simply because, were the (biological and) medical causal powers reducible to the causal powers of the physical and chemical level, the ontological gap, which I have incessantly tried to deny as holding between the two kinds under comparison, would be unavoidable. However, reductionism is a metaphysical thesis that is considered at least controversial in the current philosophy of science and a token/token supervenience approach is consistent with the argumentation I have employed. That at most token/token supervenience holds for the (biological and) medical causes vis-à-vis their underlying physical base was in fact one of the main assumptions of this thesis that I have stated from the Introduction.[321]

Nonetheless, a certain aspect of medical causation will be explored in the second part of the present thesis, in relation to the functional explanations that can be addressed to diseased organisms. More precisely, I will ask in chapter 4 how precisely the symptoms exhibited by the members of disease kinds should be interpreted, when the Humean/anti-Humean dispute is reckoned with and the biological functions are not (only) required to identify diseases, but (also) to explain them.

[321] See Sober (1999) for a discussion of the bad prospects that type-type reductionism has as far as the biology/physics interface is concerned. See also Bird (2008) for a proposal on how emergentism could in fact be viewed as a tenable position in biology.

CHAPTER 4 FUNCTIONAL EXPLANATIONS

INTRODUCTION

The present chapter continues the previous investigation into medical kinds in the direction of scientific explanations. Obviously, the pragmatic reason why classifications abound in sciences is that they smooth out the way towards explaining the behaviour of classified particulars. Simply put, pointing out the kind membership of a particular is at the same time providing an explanation for that particular's behaviour. For instance, in order to explain why a particular piece of yellow stuff conducts electricity and dissolves in a HCl/HNO_3 solution, one can indicate that the respective piece of stuff belongs to the kind gold (and gold typically conducts electricity and dissolves in *aqua regia*).

In general, natural kinds are associated with *positive* causal explanations. That is, the behaviour of a particular is reckoned with by pointing out that the properties characteristic of its natural kind are responsible for it, in a *positive* way. That kind membership can provide such positive causal explanations for the behaviours and properties displayed by various individuals is due to kinds being characterised by patterns of *determining* properties. More precisely, we have some kind specific, causal explanations because determining properties, on a basic definition, represent *positive* causes – their presence determines the emergence and maintenance of certain effects.

I have tried to show in chapters 2 and 3 that one can speak about natural kinds in somatic medicine with the same metaphysical confidence as one can talk about natural kinds in the exact sciences. Natural kinds of diseases are to be found by tracing out the patterns of co-instantiation of 'pathological' determining properties – microstructural biological properties attended by causal powers. On this account of medical kinds, what follows is that, just as in the exact sciences, the membership of an organism into a kind of disease constitutes an explanation for the pathological behaviours and properties exhibited by the organism in question.

To stick with the case discussed in the previous chapter, a valid explanation of why a

certain organism displays weight loss, goitre, tachycardia, intolerance to heat and exophthalmia is provided by indicating the membership of that organism in the GD (Graves' Disease) kind. The symptoms of such an organism are explained by its kind membership because the GD kind is characterised, on a fundamental level, by a cluster of determining properties such as increased glucose cellular absorption, high membranar concentration of the Na^+/K^+ ATPase, increased seric level of thyroxin, etc., properties which cause the emergence of those symptoms.

One can, of course, seek further explanations as to why such determining properties are in place, and as I discussed in chapter 3 (in the context of co-instantiation induction), one should be able to indicate *causal* mechanisms, either internal or external, that produce the co-instantiation by an organism of whatever determining properties are under scrutiny.

In medicine (as in the medical philosophical literature)[322] however, a large part of the explanations in use do not take an overt causal form but are rather shaped in terms of biological *dys*-functions (or failures to function). The *absence* of certain functional effects, in other words, is taken as responsible for the state of disordered organisms. The rationale for such an approach is that, since diseases are states in which biological functions fail, an explanation of why a particular organism is diseased should reveal which function (of which items) precisely fails in that organism. To return to our example, an adequate explanation in this scheme for the state of a Graves' diseased organism would be that its thyroid fails to fulfil its proper function.

This is a very broad way of speaking about the explanations founded on *dys*-functions and there are of course numerous details, depending on the account of function adopted, which need to be taken into consideration. These details will be discussed later in this chapter. However, what is important to stress at this stage is that, on this construal of pathological explanations, the 'Why' question that is being answered by pinning down *failures* of functions is intimately connected with the disease/non-disease distinction. That is because such explanations seek to explain the diseased states of organisms primarily in the sense of revealing why these are not *healthy* states. Pointing out that a

[322] See for instance Neander (1991)

certain organism has a dys-functional thyroid is *primarily* supposed to explain why that organism is not healthy, and not why it specifically has such and such symptoms.

One point I have underlined in the Introduction to this thesis was that the precise distinction between disease and health is beyond the scope of the present thesis, which is focused solely on life-threatening conditions in somatic medicine. Indeed, the dilemma of how one can decide in borderline cases which condition is a disease and which not does not influence at all the legitimacy of my enquiry in chapters 1-3 into the natural kinds of somatic medicine. In a similar vein, what I want to emphasize at the beginning of this chapter is that, when it comes to functional explanations for somatic diseases construed as natural kinds, the disease/non-disease distinction is also of little interest for the present enquiry. Here I am not concerned with the question of whether some borderline conditions should be classified as diseases or not. Rather, by focusing on organisms whose survival is threatened, I want to consider how we can reconcile the (dys-)function based register of medical explanations provided for their symptoms with the assumption that diseases are natural kinds and that the behaviour of such ill organisms should be dealt with, on an explanatory level, by indicating the positive causes characteristic of (or identifiable through) their kind.

Among other reasons, it is necessary to treat functional explanations and kind-specific explanations together because the usual (dys)-function register provides incomplete explanatory answers. As I said, most theorists would hold that pointing out that a thyroid of a particular organism exhibiting such and such Graves' Disease symptoms fails to function could in principle illuminate why that organism is diseased, as opposed to being healthy.[323] However, if the 'Why'-question is directed at the very presence of those symptoms - 'Why intolerance to heat, goitre, exophthalmia, and increased gastric transit?' - indicating in response the *failure* to function properly of a thyroid can only give an incomplete answer. Admittedly, this answer amounts to a *sort of* explanation for the symptoms at hand, because the human organism works on grounds of fine-grained interconnections between different biological items and hence, the failure to function properly of one of them explains in a sense the disruption of the whole system.

[323] There are though powerful reasons to think that functions could not draw rigorously the distinction between disease and health. See Cooper (2002) for a review of the problems that functions face in this regard.

Nevertheless, it represents a vague form of explanation, I shall argue, since it just appeals to *negative* causes of the phenomena at hand.

Informally put, what is conveyed by such an explanation is that 'something has gone wrong with the thyroid of that organism'. The 'Why' interrogation remains though, in need of a specific answer – to keep up the informal register, 'yes, something has definitely gone wrong with that thyroid and the internal mechanism of that organism has been disrupted, but still, why goitre, exophthalmia and increased transit, and not, say, shortness of breath, acromengalia and oedemas?'

Such explanations based on failures to function cannot adequately explain the symptoms exhibited by a particular organism. We need to say more than 'something has gone wrong with the biological item X' when we are dealing with an organism exhibiting specific symptoms. We can say something more, I will argue, if we make reference in our functional explanations to determining properties, which represent *positive* causes for the effects (symptoms) to be explained.[324]

The suggestion I will advance is that instead of (or complementary to) identifying the dys-functioning of biological items, we should (also) seek to identify in pathological contexts their functioning, *per se*. Kind specific explanations that refer to positive causes, I will argue, could be interpreted as functional explanations of diseases - in the paradoxical sense though in which symptoms and diseased states result from the *proper* functioning of biological items and not their functioning *failure*.

In order to show this, I shall appeal to a particular account of function and shall pursue the following rationale: to explain positively pathological symptoms is to reveal determining properties. The determining properties are micro structural characteristics of biological items attended by causal powers. For a biological item, to function is to

[324] What I shall be focusing on in this chapter are explanations connected with the relation of efficient causality between determining properties and symptoms. An equally important form of explanation is that founded on the dependence relation holding between causal mechanisms (internal or external) and the presence of the determining properties themselves. Both sorts of explanation work on the same vein - they are grounded in the causal structure of the (biological) world. The conclusions I shall draw here with respect to the former type of explanation (based on the dependency between determining properties and symptoms) could easily be applied with respect to the latter type of explanation (based on the dependency between causal mechanisms and the determining properties themselves).

produce certain effects. If functions are construed as manifestations of causal powers, and the symptoms are found to be amongst the manifestations of those causal powers, then the required causally *positive*, kind specific explanations can be interpreted as being functional explanations – explanations pointing out not the *failure*, but the *proper* functioning of biological items.

In order to forge this interpretation, I shall appeal to an anti-Humean construal of causal powers. More precisely, I shall invoke Nancy Cartwright's construal of causation, which is particularly suitable for the special sciences. I'll argue that Cartwright's *capacities* could provide the metaphysical ground for the searched-for reconciliation between functional explanations and kind specific, causally positive explanations. If organisms are viewed as *nomological machines*, I shall argue, appealing to another important notion in Cartwright's scheme, then the functioning of biological items could be invoked in order to vindicate 'pathological' explanations addressed to organisms in singular cases, that is, explanations connected with singular causal interactions involving diseases.[325]

The structure of this chapter is as follows. In §4.1 I shall briefly review the state of the art with regard to the explanatory role of functions in biology and derivatively in medicine. In §4.2 I shall discuss the notion of causal power and present its usefulness in explanations, following closely Nancy Cartwright's critique of the Humean account of causation. In §4.3 I shall present the problems that function based explanations have in the medical realm and how these problems may be solved if the functioning in pathological conditions is construed as the manifesting of capacities. The interpretation of organisms as nomological machines will be discussed in §3.4, which will insist on the causal vein of this chapter, especially on the problem of justifying singular causal claims.

[325] As far as the methodology and scope of my enquiry is concerned. I should emphasize that my purpose in the present chapter is not to adjudicate between the Humean and anti-Humean options advanced in the literature or between the various accounts of biological functions on offer, but simply to connect the discussion over functional explanations in medicine with an area in the philosophy of science that has been neglected.

§4.1 BIOLOGICAL FUNCTIONS AND THEIR ROLE IN EXPLANATIONS

What are biological (proper) functions? In what does the (proper) functioning of items consist? A preliminary but simple and useful definition is that for biological items, functioning is producing effects, and the *proper* functioning of an item is bringing forth a restricted range of 'special' effects, out of the overall effects it could and does produce. The heart, to use a stock example, produces a multitude of effects that could be considered functions: via closing and opening its valves, it produces acoustic-effects audible with the stethoscope; via endocrine-like atrial myocytes, it secrets the atrial natriuretic hormone, via its rhythmic contractions it propels blood into the arteries etc. Out of all these effects however, the propelling of blood is considered to be a 'privileged' effect, and producing it is construed as being the *proper* function of any heart.

Why is that? Because, the usual answer goes, to propel blood is *normative* and *non-accidental* for hearts. Normativity and non-accidentalness are the two crucial criteria that the 'special' effects should satisfy in order to qualify as the effects produced in the proper functioning of items.[326] They are commonly introduced in the literature using intuitive examples concerning artefacts. This is my own attempt. Think of a *clock* thrown into a window by, say, a naughty child who wants to annoy her neighbour. Even though one of the *actual* effects produced by the clock in that situation is to break the glass, intuitively, this represents an accidental effect. The clock's non-accidental effect is, in contrast, the *actual* rotation of its moving hands in order to measure time. Suppose that after the collision with the window glass, its hands do not move any more. Even if the effect of measuring time is *not* actually exhibited, its function still remains to indicate time, that is, to produce the rotation of its hands. This is what, in a modal register, it *should* do, what the normativity of its function amounts to.

What such examples involving artefacts convey then in an intuitive manner is that, out of the *actual* effects produced by an item, some are accidental and some are not. Moreover, out of the *potential* effects that *could* be produced by an item, some of them are normative, whereas others are not. Now, the intuitions attending the functioning of

[326] In the function literature, many other criteria are also discussed. Wouters (2005) for instance lists 30 criteria. However, most of them are theory laden and depend in the end on the two crucial criteria I discuss above.

artefacts apply also to a certain extent to biological items. We can grasp somehow that a heart *actually* propelling blood is producing an effect that is non-accidental, in contrast to the effect of *actually* producing noises audible with the stethoscope. On the other hand, it is somewhat intelligible that a malformed heart, which does *not* actually produce the effect of propelling blood, still *should* produce it - this is what the normativity of its function consists in.

The method of presenting the criteria of normativity and non-accidentalness by using examples of artefacts is useful (because it appeals to certain intuitions that we have) and at the same time slightly misleading (because the content of these intuitions remains implicit). What are, explicitly, the normativity and non-accidentalness of functions? How can we cash out the intuition that some of the effects produced by an item (natural or artificial) are 'special'?

The obvious thing to say would be that the proper functions address goals associated with the activity of items. In the case of artefacts, the intuitive claim that they have 'special' effects should in principle be justified by the fact that artefacts are objects *designed* for a certain *purpose* - clocks are designed for showing time and insofar as a clock remains the kind of artefact that it is, its functioning should follow the goal intended for it.

One could try to find somewhat similar justifications on the biological side. An age-old tradition sees biological items as being themselves 'designed' and their functioning is interpreted as a proof of God's existence.[327] There is an even older tradition, the Aristotelian one, according to which items have *natures* in which goals are 'ingrained', such that by functioning and producing specific goal-directed effects, items actualize their nature and hence are (or come to be) what they are.[328] Nonetheless, the appeal to Creationism or Aristotelianism has lost its appeal nowadays, to say the least - creationism in biology faded away after Darwin, and Aristotelianism was more or less expelled from science by the early modern empiricists.[329]

Yet, functions continue to be used and invoked in biology both by philosophers and

---

[327] See Ariew (2007), Bigelow and Pargetter (1990: 327, 328)
[328] See De Angelis (2003), Ariew (2002)
[329] Some Aristotelians have survived nowadays though; see for instance Megone (1998). I shall come back to the issue of Aristotelianism in the final section of this thesis with some very brief remarks.

scientists. In contemporary discussions, there are two main philosophical accounts of function that try to preserve the intuition of non accidentalness (and normativity, in the case of one of them) for the effects produced by biological items, while attempting to eliminate (or naturalise) goals.[330] I shall call them the Wright and Cummins accounts and shall describe for each of them how the 'special' effects are identified and what their (explanatory) connection to diseases is. For simplicity, I shall hence use henceforth the notation W-function and W-effects for the Wright function and the Wright special effects, and C-function and C-effects for the Cummins function and the Cummins special effects, respectively.

The Wright account, originally introduced in Wright (1973), adopts the premise that biological items have goals, but tries to naturalize them, that is, tries to justify the non-accidentalness and normativity of the functional activity of items only in terms of efficient causality, without appealing to Creationism, Aristotelian natures or any other reference to final causes. In this account, the non-accidental character of proper functions is viewed as intimately connected with the very *presence* of items in organisms. Biological items, point out the proponents of this account, were selected in the process of evolution *because of* producing certain effects - the W-effects.[331] These effects enhanced the chances of organisms' survival and consequently, the genetic bases of the items producing them was stabilised at the level of populations. Thus, since items are present in organisms *because of* producing the W-effects, it is inferred, to produce these W effects is non-accidental. Items have the *goal* of producing these effects, a goal that is revealed by the (causally efficient) process of natural selection.

Hence, in the Wright line, that hearts *actually* propel blood is a non-accidental effect – an effect following a goal – in the sense that hearts *exist* in the organisms because of producing this W effect. Normativity is also adopted. Hearts *should* produce the W - effects due to which they were selected. A genetically malformed heart *should* propel

[330] I follow here Godfrey-Smith (1993) who distinguishes between two 'families' of function accounts.

[331] Often, proponents of this account speak of the recent past (Godfrey-Smith, 1994), in order to avoid the situations in which there was a stark evolutionary change in the effects delivered by an item. There are temporal variants of the Wright line in which it is not the past evolution, but the present (Kitcher, 1993) or the future (Bigelow and Pargetter, 1990) ones that are invoked in connection to the presence of an item. More precisely, when a function attribution has as temporal context the present, the Wright effects should be responsible for the *maintenance* of *current* items; when the future is invoked, the Wright effects of an item should determine (or increase the chances of) its *future maintenance.*

blood, even if it is not capable of *actually* doing it, since the (selection related) goal of producing the W-effects is 'attached' to it in every circumstance in which it is a heart at all.[332]

'Take the heart, for instance. Biologists need and have a category that ranges across species, but hearts are morphologically diverse across different species. Some hearts have a single pump with one auricle, some a single pump with two auricles, some have a ventricle partly partitioned, and some, like us, have the two separate ventricles. Hearts are also morphologically diverse within a species, because of pathological deviations from the norm, due to disease, injury or deformity. They are all, however, *organs for pumping blood.* Not that all instances of hearts are able to pump blood. Some are too disabled. However, they are all *supposed* to pump blood; by which I mean that pumping blood is what they were selected for--it is their proper function'. (Neander 1991: 467, italics original, underlines added)

On a critical note, the main problem of this argumentation is the move from the premise that 'hearts exist at all *because* of producing certain W effects' to the conclusion that 'hearts exist *for the purpose* of producing certain W-effects'. Is this move legitimate (given that the causality involved in evolution is strictly efficient causality)? It has been argued that it is not and that goals are not naturalised but eliminated by this 'evolutionary' strategy.[333] If goals are eliminated however, then viewing the W-effects as 'special' effects of biological items appears arbitrary, and characterising them as non-accidental and normative remains just a *facon de parler.*[334]

Now, what do the W-functions explain? In biology, since the W-effects *cause* the stabilization of the genetic bases of their producing items, the W functions explain why (biological) items are present in organisms. In medicine, the scope of the explanatory usefulness of the W-functions is grounded in the alleged normativity of the W effects. More precisely, the presence of the W effects is used as a means to define the healthy states of organisms. Accordingly, the Wright *dys*-functions are taken to explain why

---

[332] For the past-focused variants of the Wright line, items *must* deliver the effect for which the item was selected. For the present-focused variants, they *must* determine the effects for which the item is currently maintained. Finally for the future-focused variants, the normative effects are those that will maintain that item.

[333] Sometimes, the premise that 'hearts exist at all *because* of producing certain W effects' is presented as being in fact synonymous to the desired statement that 'hearts exist at all *for the purpose* of producing certain W-effects'. But these two statements are obviously not equivalent, the critics of the Wright account claim. They could be equivalent only if the premise of divine creation or (or that of Aristotelian natures) was also introduced and acknowledged in the overall rationale. See for instance Schaffner (1993)

[334] See Juengst (1983: 132-135), Sandler and Agich (1995).

certain states of organisms are *not* healthy.

'Although a particular kidney might fail to filter blood, kidneys typically manage to perform their proper function, and so they standardly contribute to the fitness of organisms of the type [...] However, biological norms cannot be reduced to statistical norms, as we can see by noting that dysfunction can become widespread within a population through epidemics or major environmental disorders. A statistical definition of biological norms implies that when a trait standardly fails to perform its function, its function ceases to be its function [...] *It is the function of kidneys (both normal and abnormal) to filter wastes from blood because that is what kidneys did in ancestral organisms that caused them to be favoured by natural selection and this remains true even if renal failures become universal*' (Neander, 1991: 182, 183, italics added)

Now, let us move on to the other main approach to functions and 'special' effects. The Cummins account (1975), aims to move the focus of the discussion from the evolutionary process to the current organisation of organs and systems in organisms and to eliminate altogether goals from our construal of functions.[335] As a consequence, it rejects the claim that the functional effects of an item could be normative (in the Wright sense or on any other interpretation). However, it still preserves the claim that the functional effects are non-accidental, in a much looser sense though, to wit, in the sense that the functional effects *explain* some capacity of a relevant complex biological system in which the items producing them are integrated.[336]

A relevant complex system, according to the Cummins account, is characterised by series of fine-grained interactions between different items with a specific spatio-temporal arrangement, each item performing a specific task that contributes to the maintenance of certain capacities. The effects *actually* produced by items, which are relevant for the capacities of systems, are categorised as C-effects and producing them is deemed to be the C-function of the items in question. Since the C-effects of an item cause (in part) the capacity of the system that contains them, they explain its capacity and accordingly, are taken to be *non-accidental*.

---

[335] Variants of this account are the life-chances approach (Canfield, 1964) and the survival and reproduction approach (Boorse, 1977) *apud* Wouters (1999)

[336] See Cummins (1975), Craver (2001) and also Bigelow and Pargetter (1990: 326). Cummins presents his account as an application of an Inference to the Best Explanation strategy for biological systems. On a different note, Cummins also views the functioning of biological items as being the manifesting of *dispositions* associated to them. In fact, that the effects produced by items should be manifestations of dispositions is just another restriction imposed by Cummins in order to select his non-accidental C-effects. I shall return later to this dispositional aspect of his account in §4.3.

To propel blood is a C-function of the heart in the Cummins line because the heart is part of a complex system, the circulatory system, in which it contributes to the overall capacity to deliver oxygen, eliminate wastes, etc. This role depends on the spatial configuration of the heart in the circulatory system. It also depends on the timing of its contractions and the correlation with the different tasks of the other items in the system (e.g. the sensing of blood pressure and O_2/CO_2 concentration realized by the receptors from the aortic notch, the secreting of rennin by the renal interstitial tissue in response to Na^+ variations in the renal afferent arterioles, etc). Given that within such a system the propelling of the blood causes (in part) the overall capacity to deliver oxygen and eliminate wastes, this C-effect of the heart explains why the circulatory system is able to exercise its capacities and therefore is non-accidental.

No goal is mentioned in such a 'Cummins analysis' of the effects of an item produced within a complex system, and hence the normativity aspect of functions can be safely put aside. If in a circulatory system a heart ceases to deliver its C-effects, such that the overall capacity can no longer be maintained, this simply means, in the Cummins scheme, that the heart in question *loses* its C-function - since the circulatory system does not maintain anymore its capacity, there is nothing left to explain.

This rather loose connection between the non-accidentalness of biological effects and the explanatory role they play, together with the rejection of any form of teleology, naturalised or not, allows the Cummins theorist to pick out the C-effects of an item in a variety of systems that just have to satisfy the requirement of complexity. Within the same organism, the heart can be seen as part of the circulatory system (in virtue of its beating), as part of the renal-excretory system (in virtue of its secretion of atrial naturiuretic hormone) or even as part of a global system constituted by the whole organism. Moreover, the systems under scrutiny do not have to be localised in singular organisms. One can, for instance, perform a 'Cummins analysis' of the role of heart (as a type-organ) in the circulatory system at the level of a certain population and attribute to the heart C *typical* effects, which are non-accidental (i.e. explanatory-for Cummins) relative to organisms from the population in question.

What is particularly interesting about the Cummins functions is that they could in fact be

identified in systems with pathological capacities, part of pathological organisms. In life-threatening states of organisms, biological items are still organised in systems and their activities still complement each other and contribute to overall capacities. It is just that these capacities do not ensure the maintenance and survival of organisms, but work actively towards their demise. In an organism suffering from Graves' Disease, within the framework of the Cummins account, one can justifiably attribute to the thyroid a C-function corresponding to the C-effects that a 'Cummins analysis' of the endocrine system reveals – the increase of the number and size of mitochondria, the heightened cellular absorption of glucose and Na^+, etc. – since these are effects that contribute to the pathological capacity of the endocrine system to increase the metabolism rate, produce intolerance to heat, decrease in weight, etc.

What this means, when it comes to explanations of diseases, is that the diseased states of organisms can be paradoxically explained by the C – functioning of biological items. This explanatory use is possible at all precisely because in the Cummins account, the claim of normativity is dropped. Diseases are not identified with states in which we find C *dys*-functions of items. For a Cummins theorist, the life threatening conditions deserve a 'Cummins analysis' attributing C-effects (and thus C-functions) to items, just as much as any other states in which, say, the survival of organisms is actively maintained. When an organism becomes Graves' diseased, its thyroid simply *loses* its initial C-effects and starts delivering other C-effects. Since these new effects contribute to a newly acquired capacity of the endocrine system, one can infer that the thyroid has a new C-function explaining the diseased state of that organism. In states in which the Wright theorist could only see Wright dys-functions and violations of the normativity of the W-effects, the Cummins theorist can identify the presence of C-effects and hence attribute to items C-functions that have explanatory usefulness.

'On this view, "Wright functions" and "Cummins functions" are both effects which are distinguished by their explanatory importance. The difference is in the type of explanation. So if it is claimed, for instance, that the function of the myelin sheaths round some brain cells is to make possible the efficient conduction of signals over long distances, it may not be obvious which explanatory project is involved. This may be intended as an explanation of why the myelin is there, or it could be part of an explanation of how the brain manages to perform certain complex tasks. Sometimes the same assignment of functions will be made from both perspectives, but this does not mean the questions are the same....[Take] the case of the contribution

made by some particular mutant DNA sequence in the development of a tumor. Because the DNA sequence goes wrong in some particular way, the cancer as a whole has certain properties....On the view I have presented, we have to say that this is a case where components of the system have both Wright functions and Cummins functions, and some of the Cummins functions -- those determined by our explanatory interest in the cancer -- are opposed to the Wright functions. The Wright functions of this stretch of DNA have to do (we suppose) with regulating cell division in a particular way, which keeps the number of cells of this type at a certain level. When the mutation produces a tumor, and this tumor becomes the subject of a certain sort of investigation, the Cummins function of this bit of DNA, relative to that investigation, is a Wright *malfunction*.' (Godfrey-Smith, 1993: 7, 15-16)

Obviously, since the claim of normativity is dropped as regards the C-effects, the Cummins theorist cannot attempt to explain diseased states (in the sense of indicating why they are *not* healthy) by using the C-functions of items and claiming that items *should* perform them. This strategy, which is associated with the Wright account, is rejected. In fact, as I said, in the 'pure' Cummins account, the distinction between health and disease does not matter.

Nevertheless, some theorists whose position is close to the general Cummins line try to respond to the challenge posed by the Wright side – to wit, the challenge that a completely adequate account of function should help us in distinguishing between health and disease. Their attempt to explain diseases, as *non*-healthy states, appeal to the C *typical* effects delivered by items at the level of certain populations.[337] More precisely, diseased conditions are supposed to be explained by the lack of the C-typical effects 'normally' produced by biological items. In order to avoid the 'charge' of normativity, the lack of these C-effects in non-healthy organisms is not interpreted as a consequence of *dys*-functions (which would imply that the items involved in diseases still possess, on a modal level, their Cummins function) but simply as the *absence* of the C-typical function (in one or more biological items). However, this attempt has been severely criticised. The reason is that, in order to pick out diseases using the absence of the C-typical effects, one needs to assume that the presence of these effects is conducive to 'healthy' states, and there is hardly any way to justify non-circularly why a certain population is being chosen for delineating the searched-for C *typical* effects and not another.[338]

---

[337] See for instance McNally (2000: 309–314) Walsh (2008)
[338] See Walsh (2008: 354), Wakefield (1992), (2001)

These are some of the most important features of the current accounts of function employed in the philosophical literature. I have left out many details of the current debates. The main features I have canvassed are sufficient though for the purpose of this chapter, which is to enquire into the explanatory relevance of functions for diseases.

I shall investigate this explanatory relevance in §4.3, using the case study of a particular pathological condition, and I shall indicate there some of the problems that the function-approach has when detailed 'pathological' explanations are required for diseases and their particular instances.

These problems, as I have anticipated in the introductory part of the present chapter, have to do with the distinction between positive and negative causes and certain features of the metaphysics of causation. That is why, before discussing the particular pathological condition I have mentioned, I shall take a short detour and introduce in the next section the Humean anti-Humean debate and also Cartwright's species of anti-Humeanism, with a focus on her views about singular causation.

§4.2 CAUSAL POWERS AND EXPLANATIONS

'To explain is to reveal causes'. This motto, nowadays widely accepted, was promoted insistently a few decades ago by Salmon and the other proponents of the so-called 'causal model' of explanation, in order to emphasize their departure from the Hempelian model. What is nevertheless true is that 'to search for causes' was also indirectly Hempel's purpose, in those cases in which the laws included in the *explanans* were what he called 'laws of succession'. It is just that the causes hunted down by Hempel were Humean causes.[339] Salmon just proposed a different method to hunt them down - instead of non-accidental regularities posited beyond and above cases of singular causal interactions, he insisted on the relevance of mechanisms and transmissions of 'marks' revealed by looking directly at the singular cases. His causes were anti-Humean from this point of view and what the 'confrontation' between the two models of explanation arguably conveys is that,

[339] Hempel (1966: 53). This was also Lewis's stance; see Lewis (1986: 214-241). For discussion see (Psillos, 2002: 219), Cartwright (2002a: 5), Bigelow and Pargetter (1990: 320-323), Salmon (1998: 69)

if explanations indeed require pinning down causes, the latter task depends crucially on the metaphysics of causation one adheres to.

In this section, I shall draw first the general lines of the Humean and anti-Humean approaches to causation and then shall present in more detail a variant of the anti-Humean construal of causes developed by Nancy Cartwright, insisting upon her views on singular causation. What I shall suggest in the final part of this chapter is that non-Humean accounts of causation - and in particular, the way Nancy Cartwright develops and interprets it - solves some of the problems associated with the function-based explanations of diseases.

§4.2.1 The Humean vs. anti-Humean dispute

Suppose we have a singular causal interaction - in circumstances C, between two individuals *a* and *b* possessing the properties A and B, respectively - in which A*a* represents the cause and B*b* (or more precisely, the instantiation of B by *b*) represents the effect. How would Humeans and anti-Humeans view this singular interaction? There are significant differences among the philosophers who stand in (either of) the two sides of the debate, but the general lines of their positions could be portrayed as follows, as a preamble to Cartwright's specific views.

For an anti-Humean, we have in this case a causal interaction because there is an immanent necessity (a modal 'force' or a causal 'power') holding between the two individuals involved and their properties - A*a*, as a cause, indeed brings forth B*b* as an effect.[340] That an immanent necessity holds in this singular interaction is due to A being a *modal* property - i.e. a property whose instantiation is significant not only from the point of view of what *actually* is the case, but also from the point of view of *possibila*. There are several ways to spell out in what sense A is a property responsible for the modal 'force' of this interaction.

On a semantic level, A could be taken to be the truth-maker of a subjunctive conditional

[340] See for instance Ellis (2000: 329-330, 348-349).

with the approximate form: 'had *a* exhibited A (in the right circumstances C), *b* would have exhibited the property B'.[341] On a metaphysical level, property A can be said to be associated with the potency (or power) P,[342] in that its instantiation brings in a range of *possible* causal interactions for its bearers. More precisely, due to its association with P, its instantiation delineates a modal range of causal behaviours - in this case, the possible producing of B – that *actually* take place when the right conditions or circumstances C (the presence of stimuli, the absence of interferences, antidotes, etc.) are in place. Simply put, due to the association between A and P, A*a can* interact with particulars instantiating B. If the right circumstances C are present, it does *actually* interact.

Singular causal interactions are taken thus to spring from the potencies associated to modal properties.[343] An important consequence of this view is that any sorts of regularities involving the properties at hand represent just general systematizations of the

---

[341] One obvious connection here is with the counterfactual analysis of dispositions. For some anti-Humeans, dispositions are indeed modal properties, or features of modal properties, in the sense in which dispositional subjunctive conditionals are not just a way to set out what having a disposition represents, but to underlie the range of *possible* causal interactions in which the bearers of dispositions *can* engage in, due to their properties, as such (Bird, 1998: 233). These counterfactuals are made true by the dispositional properties at hand and not by some sort of regularity (Mumford, 1998: 87-91). What I have designated above as 'circumstances C' are reckoned with, in the counterfactual analysis, by way of mentioning (beside disposition- and manifestation-properties) also *stimulus* properties and by specifying either that the analysis holds in '*ideal* conditions' (Mumford, 1998: 87–91) in which finks and antidotes should be absent (Bird, 2007: 60) or in a certain range of situations that can be indexically circumscribed and referred to (Bird, 1998: 234).

[342] Different authors employ different terms in order to designate the causal powers and underlie certain aspects of them. 'Potency' is a term used by Bird in his (2007), which I think captures nicely the modal level of the properties in question. Cartwright's preferred term is 'capacity' (Cartwright, 1999). Other terminologies include propensity, tendency, etc.

[343] I should note that what specifically are the characteristics of modal properties, as opposed to non-modal ones, is a rather unsettled issue, which has to do with the related, problematic distinction between dispositional and categorical properties (the latter distinction having been often cashed out in terms of the relational/intrinsic dichotomy, or unobservable/observable, or non-measurable/measurable, etc.) and ultimately, with the very Humean/anti-Humean dispute - see Schaffer 2008: 4 and Ellis 2000: 329. I have assumed above the simplest approach according to which the *actual* instantiation of a modal property says something about the real of *possibilia*. For Bird, this merely means that modal properties (or potencies, in his terminology) make true certain subjunctive conditionals, *Cf.* Bird (2005: 438). For Molnar, to take another example, it means modal properties (powers, in his idiom) bear the mark of intentionality – they have a *direction* to particular manifestations, even when they are *not* exercised, as such; *Cf.* Molnar (2003) *apud* Psillos (2007: 139).

singular causal interactions resulting from capacities' manifestations.[344] Metaphysically, this means that the regularity between A_s and B_s can only be dependent on the potency P associated to A. Semantically, it means that the property A is the truth-maker not only of a specific subjunctive conditional, but also of a universally quantified material conditional that expresses the regularities in question between A_s and B_s (see below *fig.* 17).

Figure 17 **Features of the anti-Humean stance**

$\overset{P}{A}$ is the metaphysical ground for 'Aa causes Bb'
That A is associated with the potency P grounds any singular causal interaction in which B comes to be instantiated
$\overset{P}{A}\ a \overset{\text{cond. C}}{\text{---}}\!\!\text{---}\!\!>(Aa \mathrel{\square\!\!\rightarrow} Bb)$ [345]
The counterfactual analysis - the subjunctive conditional's truth maker is the modal property A.
$\overset{P}{A} \overset{\text{cond. C}}{\text{---}}\!\!\text{---}\!\!>(x)(Ax \rightarrow Bx)$
Regularities spring the stable manifestation of potencies, when the right circumstances C are present.

If we can talk about a *general* expression of the interdependence between causation, properties and powers/potencies, that general expression would have, most basically, the

[344] The direction of ontological dependency is thus from the 'bottom' level' (singular causal interactions) to the 'top' (regularities exhibited by particulars bearing the relevant properties). It is for this reason that Ellis characterises his anti-Humeanism as a 'bottom-up sort of theory, rather than top-down, [which] depends on taking a realist, non-reductionist, view of causal powers, capacities and propensities' (Ellis, 2000: 330). Ellis retains however the 'top' level of laws, just as Bird for instance. Their position amounts to what Wilson calls 'general' (as opposed to 'singularist') anti-Humeanism (*Cf.* Wilson 2005), position criticised by Mumford (2004)

[345] For a proof, see (Bird 2005:441-2). The reference to the conditions C (figured somewhat ponderous above) indicates that the regularities obtain only when the presence of stimuli, the absence of interferences, antidotes, etc. are in place. Note that the notation I have used is just meant to have a heuristic purpose. In a more fastidious notation, the conditions C could be reckoned with, firstly, by attaching a *ceteris paribus* clause in front of the quantified statement (Bird 2004: 268-9), secondly, by adding a sort of 'clause' to its antecedent, as when finks, antidotes and stimuli-properties are mentioned (Bird 2007:60) or only stimuli-properties (Bird 2005:441-2) and thirdly, by adding a 'clause' in its consequent (Ellis, 2000: 345). Note also that when and if conditions C are referred to, the least preferred solution is to bring in the *ceteris paribus* clause because, in the related context of scientific laws, the *cp* statements at hand would not be universal and would stand, evidently, for laws that admit exceptions; see Schrenk (2007)

form of a simple 'causal law' with the form 'As (can) cause Bs' - where a more precise description of the path leading from the association between A and P (A^P in my notation) to a 'general' formulation of a causal law would have to incorporate more complicated elements about the metaphysics of properties and the nature of their association to potencies/powers.[346]

Now, on the opposite, Humean side, we have no modal properties and no immanent necessity holds between *a* and *b* - *a* does not *per se* produce, *in virtue of* its possession of A, the instantiation of B by *b*. That we have, nevertheless, a singular causal interaction between the two individuals is due to a higher level regularity holding between A_s and B_s.[347] This regularity should be universal (in that it should hold no matter where and when the individuals instantiating A or B exist) and crucially, should be non-accidental (in that it is supposed to 'back up' counterfactuals that involve individuals instantiating A or B).[348]

That is, in virtue of its non-accidentalness, the regularity between A_s and B_s should constitute the truth maker of the abovementioned subjunctive conditional - 'had *a* exhibited A (in the right circumstances C), *b* would have exhibited the property B'[349] - where this truth-making relation would be just a way to express the ontological dependency of the singular causal relation between *a* and *b* on the regularity itself.

Assuming that the respective subjunctive conditional was true, the 'moderate' Humean can allow that A be attended by the potency P, but only in a loose ontological sense. Since

---

[346] In Ellis and Bird's cases, this more sophisticated 'general' expression is underlined by a metaphysics of universals whose identity is linked with their associated powers/potencies. Laws are portrayed using first order quantified, universal statements (just as the Humean laws are expressed); see Ellis (2000: 345-46), Bird (2005: 441-2)

[347] See for instance Armstrong&Heathcote (1991: 63-64) where this view is nicely summarised, and also Nagel (1961) Suppres (1970), Papineau (1991) for a more extensive exposition.

[348] For Hempel, just like for Nagel and Mackie, these non-accidental regularities are the scientific laws, where no definite indication is given as to why the laws in question back up counterfactuals; see for instance Hempel (1966) and also van Fraassen (1980) for a critique. For Lewis, the non-accidental regularities are the axioms of a deductive system which organizes all facts with the ideal combination of strength (scope of application) and simplicity (coherence of axiomatixation); *Cf.* Lewis (1986)

[349] See Armstrong and Heathcote (1991: 63-64) Cartwright (1999: 144) Bird (2005: 440). The conditions C are not usually mentioned in the classical Humean analysis and they represent in fact a great problem for the Humean strategy. One of Cartwright main arguments against this strategy, as we shall see, is that in nature we find no universal regularities and that the respective conditions C can be reckoned with only in an anti-Humean framework. See however Papineau (1991) where a response to Cartwright's contention vis-à-vis the conditions C is framed in Humean terms.

the truth-maker of that conditional is the regularity between As and Bs, what a Humean can at most permit with regard to the modal register of A is that it has its potency(ies) 'bestowed upon' it by the regularity (ies) it is engaged in.[350]

Figure 18 **Features of the Humean stance**

The regularity between As and Bs is the metaphysical ground for 'Aa causes Bb' That between As and Bs there is a constant conjunction grounds any singular causal interaction in which B comes to be instantiated.
(x)(Ax-->Bx) entails ($Aa \Box\!\!\rightarrow Bb$) Regularities are the truth-makers of subjunctive conditionals.
(x)(Ax-->Bx) entails A^{P} Potencies are 'bestowed upon' properties by the non-accidental regularities they are engaged in.

The thesis that properties could be attended by potencies only conditional on non-accidental regularities is evidently consistent with (and even required by) an approach to their nature from a non-modal point of view only. A and B are *occurent* properties, i.e. properties whose *actual* instantiation in singular situations, as such, does not provide any 'inference ticket' to a certain range of *possible* causal interactions.

The inference ticket to the realm of *possibilia* is only provided by the regularity between A_s and B_s - although properties remain occurent, even when they are considered in the context of regularities, their non-accidental conjunction is somehow able to account for the modal features of the world.[351] Put differently, A_s represent *causes* for B_s, only because

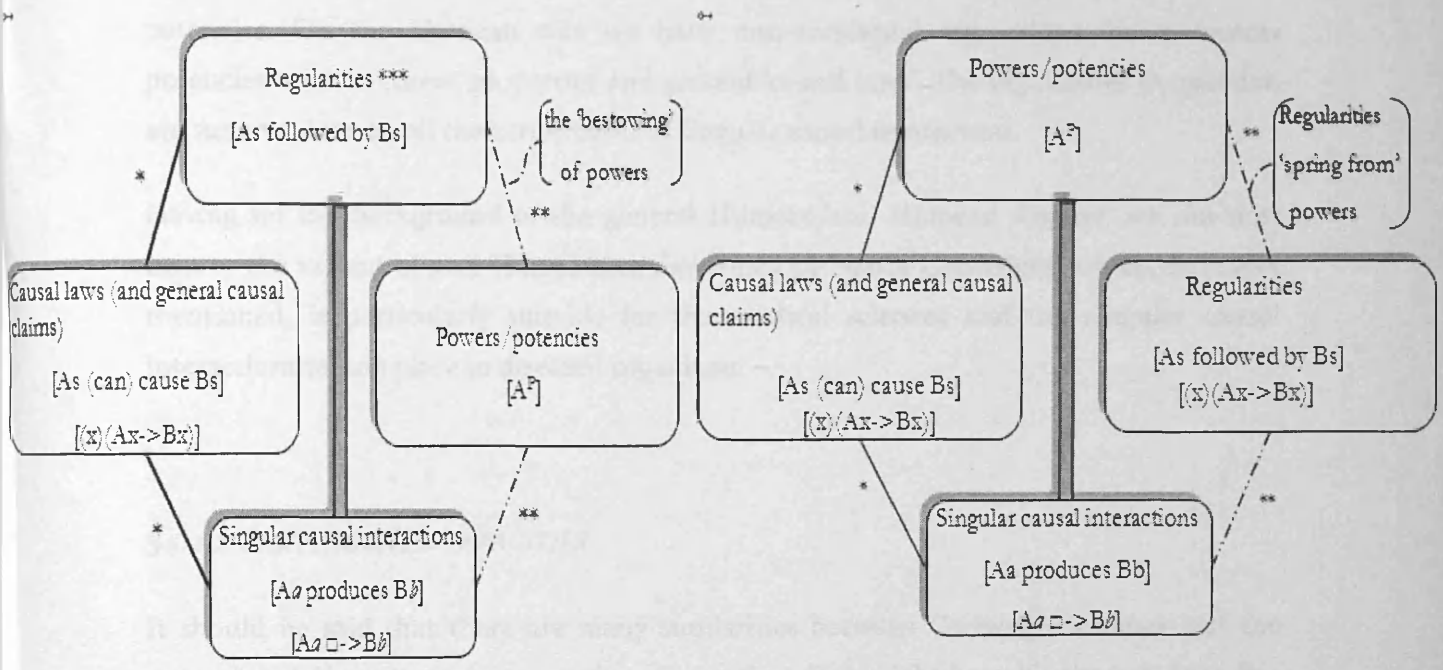
[350] See Lewis (1986:223), Hirsch(1997: 62-64)

[351] Hence the Humean Supervenience thesis espoused by Lewis - all the modal features of the world supervene on the non-modal ones; see Lewis, 1986:ix. Ellis 2000:329 gives a general characterisation of this stance on the modal features of the world, which is not limited to Lewis's views. There are classical criticisms of this position – see for instance D. Armstrong and A. Heathcote (1991), van Fraassen (1989) - that underlie a central difficulty of the Humean program: if regularities are to dissolve any modal feature of the world, why certain regularities are non-accidental cannot be justified in non-modal terms only.

A_s and B_s are part of a regularity, with some special features. What we could call as the 'causal law' that A_s cause B_s, is grounded on the regular association of the respective properties.

Two diagrams, representing what I take to be the main traits of the two opposing metaphysical schemes of causation, are presented below, in figure 19. These diagrams are simplifying and yet heuristic, I think, given the enormous literature surrounding this debate and the modest purpose of this chapter - to connect the problem of functional explanations in medicine with the discussions concerning the metaphysics of causation in the philosophy of science. The basic form of these diagrams will be reproduced several times in this section, with modified elements in its 'nodes', depending upon the context of the discussion.

***Figure* 19 The Humean (left) vs. anti-Humean (right) rationale on the relation between regularities, capacities, causal laws, general and singular causal claims. The connecting lines indicate ontological, epistemic and semantic dependencies.**



* Regularities [As followed by Bs] ground general causal claims [As (can) cause Bs] and (thus) ground singular causal interactions [A*a* produces B*b*]. The causal laws corresponding to the general causal claims are formulated using universally quantified statements [(x)(Ax->Bx)].

** Regularities [As followed by Bs] ground potency attributions [A^P] and (thus) ground counterfactual scenarios [A*a* □->Bb]. Potencies (or powers) are 'bestowed upon' properties by regularities.

*** The regularities in the top box are the brute, basic regularities, out of the systematisation of which Lewis for instance derives his Best System theorems – the Humean laws, as non-accidental regularities. Only a part of them are causal, and the latter are figured in the left-hand box.

* Potencies associated to properties [A^P] ground general causal claims [As (can) cause Bs] and (thus) ground singular causal interactions [A*a* produces B*b*]. The causal laws corresponding to the general causal claims might as well be formulated using universally quantified statements [(x)(Ax->Bx)]. These statements might also contain a clause referring to the conditions C in which the manifestations of capacities obtain. Most often, the respective clause is added in such a way that the statements remain universal and the corresponding laws strict.

** Potencies associated to properties [A^P] ground regularities [As followed by Bs] and (thus) ground, as a *facon de parler*, counterfactual scenarios [A*a* □->B*b*]. Note that the regularities in question might be the brute, basic regularities of the naïve Humean or the sophisticated regularities *a la* Lewis [(x)(Ax->Bx)]. At any rate, these regularities are taken to 'spring from' the systematic manifestation of capacities.

So, on the anti-Humean side we have immanent necessity, potencies attending modal properties which are manifested in actual singular causal interactions and regularities which represent simple systematizations of the singular cases of interactions. The 'causal laws' express in a 'general' way the nature of the association between properties and potencies. On the Humean side we have non-accidental regularities which bestow potencies upon occurent properties and ground 'causal laws'. The regularities in question are supervenient on all the actual cases of singular causal interaction.

Having set the background of the general Humean/anti-Humean dispute, we can now turn to the variant of anti-Humeanism developed by Nancy Cartwright, which, as I have mentioned, is particularly suitable for the medical sciences and the singular causal interactions taking place in diseased organisms.

§4.2.2 *CARTWRIGHT'S CAPACITIES*

It should be said that there are many similarities between Cartwright's views and the general anti-Humean program, perhaps more than Cartwright herself acknowledges. For instance, Cartwright accepts (and was actually one of the very first proponents of) the anti-Humean account of regularities according to which they 'spring' from the 'potencies' (her preferred term being *'capacities'*) associated with properties. Regularities *just* systematize singular causation cases that take place when capacities manifest themselves in certain circumstances, argues Cartwright.[352] There are, however, some differences in approach (or at least some differences of accent) and some of them will be important for the medical cases I shall be looking at in the next sections.[353]

---

[352] See for instance Cartwright (2007a: 8). Menzies also proposed that capacities could be taken as truth-makers of counterfactuals (see Menzies 2002), a move rejected by Cartwright, however, for un-specified reasons (Cartwright, 2002c). In her (2007a) such a counterfactual-based feature is I think adopted tacitly though - see (2007a: 24-5) and the discussed difference between the *exercise* and the *manifestation* of a capacity.

[353] It should be noted that Cartwright is not always clear about the precise metaphysics that underlies her account of causation; for instance the bearers of capacities are sometimes considered to be properties, other times natural kinds or systems; see Drewery (2001) and Paul (2002) who note this lack of clarity. My employment of the category of property in the above presentation of her views is in a sense just a way of cutting the Gordian knot.

One point of difference is that capacities can be *exacerbated* or *diminished* by specific stimuli, specific 'causal contexts', and this aspect will be crucial for my analysis of medical causes in §4.4.[354] Another difference is that the regularities that can be derived from capacities are not spatio-temporally unrestricted[355] - these regularities may only obtain in specific conditions. The latter is a particularly important aspect of Cartwright's argumentation because, *inter alia*, it extends the scope of the capacities-rationale to the special sciences, where the lack of universal regularities is notorious.

In the mainstream anti-Humean interpretation, the resulting regularities are usually expressed as (first order) universally quantified statements, which have at most a clause introduced in the antecedent or consequent of the respective material conditionals, in order to indicate a set of circumstances C influencing the manifesting of capacities (the presence of proper stimuli, the absence of interfering factors, antidotes, etc). For Cartwright however, capacities give rise to *local* regularities, circumscribed precisely to those sites in which the clause delineating the circumstances or context C is satisfied. In other words, Cartwright prefers to keep in the antecedent of the respective conditionals only the (bearers of the) capacities and in the consequent only the (bearers of the) manifestations, while all the other features enabling the manifesting of capacities are exiled in the quantification area. In the notation I have used in the previous section, in Cartwright's view the attribution of capacities or potencies to properties can be used to derive only *local* regularities in that A^P entails (for all x in the circumstances C) (Ax-->Bx).

The settings in which the conditions C are present are called 'nomological machines'. The latter are certain 'arrangements' of factors that enable the full manifestation of potencies by 'shielding' them and their bearing individuals from any interference. Experimental set-ups provide the classic examples for such nomological machines. Now, Cartwright stresses that there is an obvious distance between what happens in the contexts resembling the 'inside' and the 'outside' of nomological machines. *'Inside'* a nomological machine, the capacities are enabled to manifest such that the behaviour of their bearers can be expressed in the form of regularities - any individual possessing A interacts with individuals possessing B. Thus, the regularity between A_s and B_s, does correspond to how

[354] Cartwright (1989) (1995: 154)
[355] Cartwright (1999: 25-27, 148)

the individuals possessing the properties in question *actually* behave, i.e. to what happens on an occurent level.

'*Outside*' of nomological machines, the capacities are still present, but stimuli are modified, interferences show up, etc. and the behaviour of the capacities' bearers cannot (justifiably) be described in the form of regularities. That is to say, when the 'shielding' disappears, the statement describing the regularity between A_s and B_s just stands for *in abstracto* correlations between properties with no direct connection with how the individuals which instantiate the properties in question actually behave.[356] What the presence, as such, of capacities guarantees to us are (just) 'causal laws' of the form 'A_s have the *capacity* C' 'A_s *can* produce B_s' 'A_s *tend* to produce B_s', etc.[357] These are, one could say, spatio-temporally unrestricted, universal claims, but on a *modal* level – they point to *possible* causal behaviours. In other words, they are *not* regularity-like claims about what *actually* is the case. As for what *actually* happens in an 'outside' context - depending on the stimuli/interferences in a particular situation, the manifestations of a capacity can be exacerbated, diminished or eliminated.[358]

So, the distance between the 'inside' of nomological machines and the 'outside' is the distance between two different sorts of causal contexts, which matter for our general and singular causal claims. When we deal with causal contexts that are 'inside' - that is, causal contexts in which we have a specific 'shielding', stimuli and a certain arrangement of causal factors -the regularities constitute an inference ticket about what, on an *actual*-ist level, takes place in singular causal interactions. When we deal with a causal context that can be described as 'outside' a nomological machine, we know that capacities are still present,[359] but this only offers us an inference ticket as to what, on a *modal* level, *could* take place in singular causal interactions. We need in addition *local* information, information about stimuli, on how they change, what interferences show up, etc, in order

---

[356] Cartwright (1999: 65-66,71,139-143)
[357] Cartwright (1999: 59, 66, 138) (1989: 142-145, 154)
[358] '[capacities-claims]1. are claims about what systems with the cited features...can cause 2. the effect described may not occur but only 'contribute' in some appropriate way to the actual outcome in any given case 3. the effect or even the successful 'contribution' will only occur ceteris paribus...[the] prescriptions for filling in the ceteris paribus clauses...involve a whole range of categories from the ontology of capacities that are abhorrent to the Humean: triggering, shielding, enhancing, interfering and the like' (Cartwright, 1995: 154).
[359] Cartwright, (2002b: 3-4)

to infer what type of manifestations are *actually* present.[360]

'[there are] three levels of causal claims important in scientific explanation, prediction and application: singular claims, causal laws and ascriptions of capacities. Capacities claims, I said, summarize a range of outcomes a subject *can* cause. Causal laws are more specific: They describe, usually probabilistically, what causal relations actually obtain in a specified kind of situation. *Causal laws have thus the advantage over capacity claims that they can be used in a mechanical way for making predictions and also that they can be established by induction. But in both cases this is true only as long as the target situation and the situation to which the laws are relativized are the same. Their strengths are thus the flip side of their weaknesses. They describe what really does happen when the situation is so and so.* And they are proper laws: they describe what would happen, were the situation like that. But by their very nature they do not describe what would [actually] *happen* were the situation different' (Cartwright 1995: 154, italics added) [361]

The relation between 'inference tickets', capacities and contexts that resemble the 'inside' or the 'outside' of a nomological machine can be usefully figured as follows.

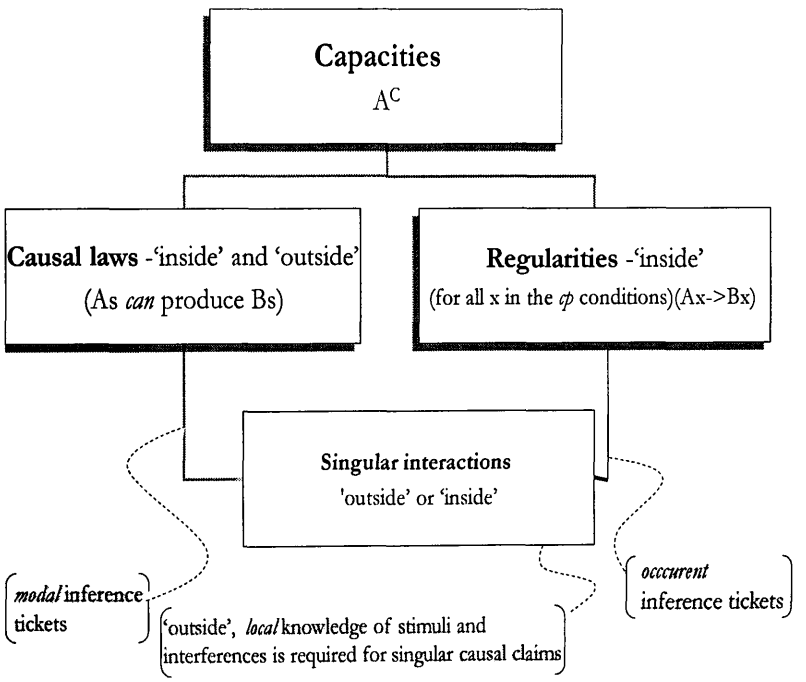


***Figure* 20 Capacities and inference tickets**

'*Inside*' nomological machines: A^C entails (for all x in the *cp* conditions)(Ax->Bx), where the *ceteris paribus* clause indicates the '*inside*' of nomological machines. [C 1999:139-143]

'*Outside*' nomological machines we have (As *can* produce Bs), or (As *tend* to produce Bs) or (As *have the capacity* to produce Bs) [C, 1989:142-145]

The regularities of nomological machines – inference tickets, on an *actual*-ist, *occurent* level, for what happens in causal contexts resembling the 'inside' of nomological machine.

[360] See Cartwright (2005: 2) and the discussion of 'background knowledge.

[361] The 'causal laws' that Cartwright refers to in this quoted passage are the regularities from the 'inside' of nomological machines. In a more recent paper (Cartwright, 2005) as 'causal laws are designated the 'capacity claims' of the form 'As *can* cause Bs' and I shall primarily refer to this more recent use in the main text. See also Cartwright (2002a: 12-16), (2007a: 58)

Now, that there are no universal regularities is in fact the main argument employed by Cartwright against the Humean scheme. Cartwright complains that the Humean ignores this possible change in the causal *context* and thus fails to distinguish between, as it were, the 'internal validity' of our *general* causal claims and their 'external validity'.[362]

Problems concerning the vindication of *singular* causal claims evidently ensue for Humeans. Since they take their regularities as universal, they construe the 'causal laws' grounded in them as having unrestricted scope and thus capable of vindicating any singular causal claim. However, both in the case of deterministic and probabilistic correlations, insists again Cartwright, we have no reason to think that the corresponding causal claims formulated by the Humean will be universal. They have a scope limited to the context in which they were assessed, and this applies both to the exact and special sciences, irrespective of the method used to ascertain causal claims - in laboratory settings or at a population level.[363]

The general causal claim that, without exception, A_s produce B_s in the context C (or A_s produce B_s with such and such probabilities in the causal context C) could not be taken *universally*, in the form that A_s produce B_s *simpliciter*, across all causal contexts (or that A_s produce B_s with such and such probability, over any causal context, over any population, etc).[364] Hence, for any putative causal interaction located in a causal context different from the one in which regularities were ascertained, what causes what cannot be ascertained by the Humean 'causal laws'.

An illustrative 'case-study' is Cartwright's discussion of Newton's second law. This law can be read in a Humean key as expressing a regularity. In order for this law to be applied, tested, used in inference, etc., says Cartwright, one needs certain bridge principles, that

---

[362] Cartwright (2007b: 220)

[363] When discussing exceptionless, deterministic regularities (which are usually referred to in connection with 'experimental settings'), Cartwright has in mind both the regularities of classical physics for instance, and at the same time the regularities established on a microstructural level in the special sciences - see Cartwright (1999). When discussion about the probabilistic regularities, Cartwright usually refers to the population-level assessments from the special sciences; see Cartwright (1989).

[364] Cartwright (1989: 142) (1999: 24-27) (2007a: 58)

are in fact provided by the special force laws.[365] Admittedly, the special force laws allow the application Newton's second law *as* a regularity, but at the same time restrict its scope. They bring in an explicit clause specifying the quantities at work, but also a tacit clause forbidding any other sort of influence to be exercised over the experimental setting/context in which the regularity is tested.[366]

> Consider Newton's second law, $F = ma$. What does it say? Many...take it to describe a strict regularity. I think that it does so only conditionally. The claim we are entitled to believe from the vast evidence in its favor is this: *if nothing that affects the motion operates that cannot be represented as a force,* then....The bridge principles match some theoretical concepts with concepts 'antecedently understood'. In the case of Newtonian mechanics the primary bridge principles are the special force laws. These license a particular theoretical description – e.g. '... is subject to a force $F = GMm/r2$' or '... is subject to a force $F = eoq1q2/r2$' – given a description in the vocabulary of masses, distances, times and charges– e.g. '... is a mass m located at distance r from a mass M, or '... is a charge $q1$ located at distance r from charge $q2$'. Bridge principles provide strong constraints. The theoretical descriptions – in our example the individual force functions from the special force laws – are allowed *only if* the corresponding descriptions in 'antecedently understood' terms are satisfied. (For example, "The force on a mass of size m is $GMm/r2$ if and *only if* m is a distance r from a body of mass M".)...It is because of the issue of evidence that I urge that the bridge principles of these theories are so strongly constraining .....Can all causes of motion be correctly described using just the descriptions that appear in the bridge principles of Newtonian mechanics? To all appearances, not. We have millions of examples of motions that we do not know how to describe in this way....We have succeeded in applying Newton's second law to a vast, vast number of cases – but always of the same kinds: the ones that appear in our bridge principles. (Cartwright 2007a: 36-37)

The two sorts of clauses, the tacit and the explicit one, amount in fact, says Cartwright, to a restriction of the scope of the regularity that the respective 'bridge principles' are supposed to 'translate'. This is salient if we look at a singular causal situation – for instance a banknote dropped from the steeple of a cathedral, in a context in which, beside gravitational attraction, other forces are also at work. As Giere explains, there are no

---

[365] Even if Cartwright does not put it in this way, the special force laws could be read as 'causal laws' that can be derived from the regularity of Newton's second law. This reading is in a sense inadequate, since a Humean would want to view the special force laws as expressing regularities (and *derived* causal claims) just as much as he or she would want Newton's second law to express regularities. But what the special force laws bring about are more 'shaped' regularities, that introduce specific contexts, specific properties, etc. I would suggest this reading for the purpose of symmetry with the case of probabilistic regularities, where from frequencies are derived 'causal laws'. In addition, this reading would make more intelligible the two diagrams introduced in figure 20, which are, admittedly, more suitable for special sciences.

[366] Evidently, Cartwright underlies at this point what the two clauses are in fact meant to delineate – namely nomological machines, 'inside' of which capacities are enabled to manifest themselves and thus 'spawn' regularities; see Cartwright (1999: 50).

specific models suiting the drifting of such a banknote. In other words, once the tacit clause precluding the existence of interferences is violated in a particular context, the general causal claim represented by Newton's law cannot satisfactorily deal with certain singular causal interactions.

'Imagine dropping a banknote from the top of St. Stephen's Cathedral (in Vienna) and watching it drift down to a spot in the square below. Within the context of classical (non-relativistic, non-quantum) physics, one might suppose that, in principle, the motions of the banknote could be fully accounted for in terms of Newtonian forces. This supposition Cartwright...rejects. She calls adherence to such a supposition (scientific) fundamentalism. Fundamentalism, in this context, amounts to generalizing the dominant science of the day to cover literally everything. Cartwright insists that there is no justification for this wild extrapolation from such well-known cases as dropping a cannon ball from the Leaning Tower of Pisa. In the case of the banknote, the effects of air-resistance and random gusts of wind are comparable to those of gravity. There are no specific models in mechanics, or in hydrodynamics, that fit this case. The falling cannon ball, by contrast, is an example of what Cartwright calls a *nomological machine'* (Giere, 2000: 527)

A falling cannon ball has a motion 'governed' by the regularity of Newton's second law because its (natural) context is similar to that of a nomological machine - except the earth's attraction, all other causal influences are negligible. The cannon ball is in a sense 'inside' a natural nomological machine. What happens 'outside' with a banknote? To say that it is still subject to the same (non-accidental, 'governing') regularity is a matter of faith, or fundamentalism, stresses Cartwright, in a rather picaresque way. [367]

Now, Cartwright makes the same point with regard to the population level, imperfect regularities out of which the Humean wants to extract probabilistic 'causal laws' - the causal laws in question could not have universal bearing. The imperfect regularities from the population level result in one way or another from an application of Reichenbach's principle - any correlation indicates causation, unless there is another *cause* that 'screens

[367] Cartwright (1999: 27). The analysis of the microstructural assessments performed in laboratory conditions proceeds over the same path as the analysis of Newton's second law. Whereas in the case of Newton's second law we have a natural nomological machine, the laboratory conditions set out artificial nomological machines, which have the same advantages and tribulations as the natural ones; see Cartwright (1999: 60). Importantly, Cartwright warns against the semantics of 'mechanism', a notion that seems directly linked to microstructural assessments but whose use could be misleading, because its meaning possesses a definite anti-Humean register. Mechanisms are, after all, 'shielded' settings. The Humean proponent might make use inappropriately of this notion and Cartwright is keen to point out that, if Humean mechanisms amount to (further) microstructural regularities, then the problems issuing from the variability of causal contexts remain the same; *Cf.* Cartwright (1999: 138-139, 143).

off' the correlation. To keep with our simple notation, any imperfect regularity between A and B in a given population - pointed out by an increase in probability (Pr (B/A) > Pr (B/¬A)) - underlies a probabilistic causal law, unless, by taking into account any factor F from the population in question, the regularity disappears. [368] In other words, Reichenbach's principle says that A_s cause B_s, if the presence of any (other causes) F_s of B_s makes no difference to the rise in probability - Pr (B/A.F) > Pr (B/¬A.F).[369]

The Humean needs however to read Reichenbach's principle in a reductionist way. The regularities grounding the probabilistic causal relation between A and B - which are not screened off by the other *cause* C present in a test populations - should constitute a probabilistic pattern in which should play a part not only A and B, but also **F** itself. F is the factor whose presence, even if not making a difference to the probabilistic dependence between A_s and B_s, plays the role of a background, or causal context in the final expression of the rise of probability [Pr (B/A.F) > Pr (B/¬A.F)].

If a Humean is to stick to his principles (i.e. to the view of properties as occurent and to the dissolving of causation into regularities), claims Cartwright, then the probabilistic pattern between A_s and B_s established in a test population could not be projected at the level of another population in which we would have a different background/context. For a different population in which the present factors were A, B and F*, we would have no reason to expect an increase in the probability of B given A (irrespective of whether we conditioned on F* or not). That is, a Humean who stuck to his principles would have to submit that the 'causal laws' that A_s cause B_s could not be simply deemed universal. It holds in the test population T, given a certain causal background in T - who knows what might happen elsewhere, in a different population?

This is for instance Cartwright's argumentation against the Geiger/Suppes approach to probabilities, presented in her (1989) – the volume in which she actually introduced the notion of capacity. According to the Suppes approach, the rise in probability that should

[368] I shall be using 'A', 'B' (and 'F') as standing either for a property or a causal factor (the causal factor that instantiates the property in question) depending on the context of the discussion.

[369] This principle is to be found in all population-level techniques of 'hunting down' causes. In the causal path modelling, it is expressed in the condition that the error terms of the recursive equations are independent, in Bayesian nets by the Markov condition, in the Geiger causality by the condition that populations should be stratified such that all the other causes of the effect in question should be present or absent; see Cartwright (1989), (2007b)

ground the probabilistic 'causal law' linking A and B in a test population, should be 'partial conditional' - i.e. established by 'stratifying' the population under study in 'homogenous' subpopulations, i.e. in subpopulations in which all the other causes of B were either present or absent. Hence, for a test population T in which all the other factors connected to the effect B were $\{F_1,..., F_t\}$, in order to conclude that A is a cause of B, it should be the case that: $Pr_T (B//A+/-F_1+/-F_2....+/-F_t) > Pr_T (B//\sim A+/-F_1+/-F_2....+/-F_t)$.

The latter expression of probabilistic dependencies is called 'the CC formula' for a test population T.[370] Importantly, the formula has a stronger, universal corollary saying that beside the test population T, in *any* other population I - stratified with respect to all the other causes of B $\{F_1,...F_i\}$ - we should expect that $Pr_I (B//A+/-F_1+/-F_2....+/-F_i) > Pr_T (B//\sim A+/-F_1+/-F_2....+/-F_i)$. Cartwright argues that the Humean cannot justify precisely this *universal* form of the CC principle, requiring the probabilistic causal claims established in a population T to hold in all the other populations.

If A is viewed in an occurent manner – i.e. reductively taken as a part of some regularities, and not as a *cause, per se* – then the change of background when moving from the test population T to another population I, makes the Humean regularities in T irrelevant for

---

[370] There are two possible interpretations of the CC formula. One position, the moderate one, is that the CC formula amounts to a *methodology* to track down causes in test populations, which has no bearing on the basic Humean metaphysical point - even though 'all the other causes' need to be introduced in the methodology because of our epistemic limitations, 'if we knew all the statistical associations [imperfect regularities] in the universe, there could not, in addition to these, be further facts that we had somewhat failed to mention [in our methodology of tracking down causes]' (Cartwright, Dupré 1988: 522). *Contra* this moderate Humean position, Cartwright argues that the causal claims established in a population do not (or cannot justifiably be said to) have universal bearing. I present in the main text this argument. The other Humean (radical) position is that the CC principle does amount to a reduction of causation to probabilities since 'all the other causes' that need to be conditioned on in the process of 'stratifying' a test population, can simply be picked out in a reductive way as the factors that are probabilistically related to the effect in question. Cartwright argues in her (1988) and (1989) that this radical position is untenable and the criticisms are reiterated in her comments against the assumptions of the Bayesian nets techniques, in Carwright (1999). In short, Cartwright's point is that, for probabilistic causes, we might have correlations between factors, which are not screened off by anything and which yet do not underlie causal relations. Given a probabilistic common cause producing an effect and a side effect via a certain pattern of 'firing', the effect and the side effect could still be correlated, even after reckoning with the common cause and calculating the partial conditional probability. That is to say, not all probabilities dependencies indicate causal relations (see Cartwright's comments against the Markov condition used in Bayesian nets, in her 1999: 107-111). On the other hand, probabilistic independencies do not entail that no causal relations are in place; that is, not all causal relations manifest as probabilistic dependencies (against the Faithfulness condition in Bayesian nets techniques; see her 1999: 117-118). A factor (cause) might produce an effect via two pathways, an enhancing and a preventative one, such that the overall influence of one over the other is cancelled and appears invisible in the statistical evaluation. The case discussed by Cartwright in her (1989) is Hesslow's birth control pills, with their dual effects on thrombosis.

causal inference.

'The [CC] formula says that C causes E if and only if E is greater with C than without C in every homogenous background ...why expect any pattern of probability relations from one background to another, let alone the totally uniform pattern implied by the universal quantifier?...what is supposed to follow from one increase in probability of E on C in one of the specially related populations (test populations)?...what follows is that, in those special test situations, it can be regularly relied on that some Cs will produce Es...The point I want to make about this argument is that it justifies *a very local* kind of causal claim: if in a given test population we see the increase in probability that we are looking for, that guarantees that Cs cause Es *there in that population*. But it does not tell us any more' (Cartwright 1989: 142-145, italics original, underline added)

Of course, Cartwright is keen to point out that making causal inferences across different backgrounds/contexts requires giving up the occurent construal of properties/factors and introducing capacities.[371] That we can expect As to cause Bs not only in T but in all populations where all the other causes are conditioned on, implies that As have the *capacity* to produce Bs, and this gives us a modal inference ticket from T to all the other populations.

This sort of argumentation is also presented by Cartwright against the other methods of ascertaining probabilities at the level of populations.[372] The purpose is the same –to show that the (Humean) causal laws can only have a narrow, non-universal scope - again, given that A_s and B_s in a population, one cannot exclude that in another population A_s have no influence upon B_s, or might even prevent B_s. The appealed to examples are also variegated. For instance, skill loss during unemployment *might not* perpetuate high levels

[371] *Cf.* Cartwright (1989: 142-145). Notably, the moderate Humean could retort here that 'capacities' could be introduced just for methodological reasons, because of our epistemic limitation. That is, the Humean could claim that, if we knew all the regularities in the universe, 'capacities', the alleged immanent necessity attending them and the modal inference ticket across different contexts would dissolve into 'a kind of meta-pattern, a kind of recurring structure in first order patterns of association' (Papineau 1991: 409). See also Morrison (1995) and Psillos (2008) for a critique of the same type, which tries to 'dissolve' capacities into regularities and see the dependence on the 'causal context' of causal manifestations as just 'more grist to the Humean mill' (Papineau, 1991: 409).

[372] The whole argumentation concerning probabilistic causes (and the need to introduce capacities) can be applied for deterministic causes, and Cartwright does provide such an application, using the recursive equations of causal path analysis. Instead of enhanced probabilities we have enhancements of the strength of effects, evaluated by the coefficients of right-sided 'exogenous' variables, when all the other causes are being 'fixed'. See Cartwright (1989), and also Irzic (1992) for a critical discussion. The Bayesian nets are criticised in Cartwright's (1999) and (2007b).

of unemployment in spite of the maximizing utility 'causal law',[373] or, to take a different situation, in spite of the well documented 'causal law' that smoking causes lung cancer, one cannot exclude that there might be a population in which smoking would prevent this cancer.[374] The latter is obviously an example from the medical sphere.

When discussing cases of singular causation in which the lack of universality for the causal laws shows its consequences, it also happens that Cartwright appeals to examples connected with the medical realm.

"The more usual way to use probabilities to treat singe cases is in a two-step process. We use the probabilities to establish generic causal claims, then – like Donald Davidson – we insist that all admissible singular claims fall under a generic law. So i) if we take increase in probability as a necessary condition for a generic causal truth, "*C*'s cause *E*'s" and ii) we take as a necessary condition for "this *C* caused this *E*" that it be generically true that *C*'s cause *E*'s, then we would have as a result that increase in probability of *E* on *C* is a necessary condition for the related singular causal claim. I have already explained why I do not accept the first step. I am also suspicious of the second, whether or not we accept the first. I particularly worry about the second in medical and psychiatric settings.... Let us begin anecdotally. I offer Emily a regime of diet to lose weight. She says, "That kind of crash diet never works". I reply, "It works for me". And I do so with good reason. I am fairly careful about these matters. I have tried this diet frequently. It always works. I know about the possibility of spontaneous weight loss and about the placebo effect, about long-term vs short term outcomes *etc.* and I have evidence these are not a problem. Now there *may* be some description, *D*, of me that fixes *in a law-like way* the efficacy of this diet for people who satisfy *D*. In that case we would have a reference class, picked out in a non question-begging way, in which *P*(weight loss/diet) > *P*(weight loss/no diet). But there may not be such a description. There may simply be individual variation. To insist that there is always such a description is to let a big – and insecure – metaphysical assumption guide our methodology. That I think is a wrong thing to do." (Cartwright (2005), italics original, underlines added)

There might be then *individual* variation of the causal contexts, which is manifest in the particular cases of individual organisms, such that the general 'causal laws' are incapable of vindicating singular causal claims. This brings us close to the next part of this section, in that there is a straightforward application of Cartwright's criticism of the Humean approach on causation (and her urging for the introduction of capacities) to medical cases and the problem of medical diagnoses.

Before moving on to the medical side of the discussion, a final note should be added, with

[373] Cartwright (2007a: 73)
[374] Cartwright &Dupré (1988: 526-27)

respect to our prospects of hunting down singular interactions and the epistemology of causation that Cartwright employs. It might appear, given her critique of Humean regularities and even her own framework of capacities – which provides, as we have seen, two sorts of inference tickets, an occurent and a modal one – that our prospects of advancing warranted singular causal claims are rather bleak. The Humean 'causal laws' have a limited scope. Likewise, the occurent inference ticket offered by capacities is reduced to certain special contexts. For any interactions that occur in 'outside' contexts, we only have left the modal inference ticket, which does not say much about what *actually* happens, or should be expected to happen, in singular situations. Aspirins *can* bring relief to headaches, inverted populations of molecules *can* produce lasing, positively charged particles *can* attract negatively charged ones, etc. These are modal claims, which are consistent with different manifestations showing up on an *occurent* level, if the variation in the causal context is great enough, and which could even stir questions about the existence and individualisation of capacities.[375]

The picture is not that bleak however. The framework provided by capacities has a 'constructive' part, precisely because the 'causal context' can sometimes by identified or controlled. This has two consequences. First, regularities can be created. Controlling a certain causal context ensures that the manifestation of a single capacity of a causal factor happens regularly. As I said at the beginning of this section, Cartwright calls such arrangements in which 'ideal' causal contexts exist 'nomological machines' and in general, identifies them with the highly contrived settings of laboratories. For all the singular situations in which the causal context of a nomological machine is present, we can univocally identify the causal role (the *manifested* capacity) of a certain factor and predict the actual outcome of the associated capacity.

Second, in singular situations in which no regularities are at hand, given that we know that the capacity is there and tends to manifest, we can identify the causal context that influences its manifestation. The rationale here is rather different from the one used in the case of nomological machines. Whereas in the case of nomological machines, where regularities have to be created, one has to preserve the triggering factors (if triggering factors are needed) and eliminate *all* interferences, in the singular cases where

[375] Cartwright raises herself some of these questions in Cartwright and Dupré (1988).

regularities are absent one has to assume the existence of capacities and try to identify all triggering factors *and* interferences. This allows one to determine whether the effect is caused or not, and if caused, whether in a diminished or enhanced way.

These two consequences of controlling and identifying 'the causal context' are interrelated. The first consequence, simply put, allows us to track down capacities. Even if, on a *metaphysical* level, capacities do not depend on regularities (as all anti-Humeans agree), the regularities of the nomological machines represent the right starting point, from an *epistemic* point of view, to identify capacities and their manifestations. The second consequence allows us to use knowledge of capacities in that we can detect the manifestations of these capacities, outside nomological machines, if we also have contextual, information of whether interferences act as exacerbating, diminishing or 'antidote-like' stimuli.[376]

These consequences will play an important role in the following section, where I shall assume Cartwright's framework and shall try to apply it to the issue of pathological functional explanations, by way of asking what capacities could be involved in diseases. For now though, let us turn to the application in medical cases of Cartwright's criticism of Humeanism.

§4.2.3 *Cartwright's rationale applied to medical causes*

In the medical cases, we have two classical ways of ascertaining causation – the epidemiological study of frequencies at a population-level and the physiopathological assessments in laboratories at a microstructural level.

These assessments are used to infer 'causal laws' or general causal claims. The latter are either probabilistic (e.g. smoking causes cancer with such and such probabilities) or deterministic, exceptionless (e.g. aspirin ingestion over 5g/d induces an inflammatory

[376] See Cartwright's discussion of 'background knowledge in her (2005:2) and the comments in her (1999) on the information needed to built up experimental settings. The respective information '*is not itself a report of any regular association* – neither a real regular association (one that *occurs*) nor a counterfactual association (one that might occur)' (Cartwright 1999: 143). See also her discussion of 'rules of thumb' in Cartwright (2007a:62-64)

response in the renal interstitial tissues), depending on the sort of evidence drawn upon. Usually, the epidemiological evidence is a basis for probabilistic causal laws whereas the exceptionless ones come from laboratory, microstructural assessments.[377] Of course, we have in addition medical diagnoses (e.g. this case of pulmonary cancer was caused by smoking, or, this case of renal failure was caused by aspirin intoxication) which use the general 'causal laws' to draw singular causal conclusions about particular organisms.

Precisely how 'causal laws' are inferred from the evidence is a different discussion to which I shall return in §4.4. What is important to note at this point is that the medical literature on causation is largely agnostic about the Humean/anti-Humean dispute. In other words, insofar as the explicit medical discourse is concerned, the Humean is free to read it in a reductionist way - as amounting to causal claims that have regularities as their truth makers.[378]

Now, there is a natural extension of Cartwright's anti-Humean critique to the medical realm, as the realm of a special science. Cartwright mainly analyses case studies from physics and economics but also gives some examples from epidemiology, which enable us to see why the Humean might be wrong when failing to limit his/her causal claims, given the variation in the causal contexts that occur in medical situations.

For instance, as I have mentioned in the previous chapter, Cartwright holds that in spite of the 'causal law' holding that smoking causes cancer with such and such probability, we might have a population in which smoking was in fact beneficial for the lungs.[379] Not many examples are given for microstructural assessments,[380] but I take it that, since 'anything can cause anything else',[381] we cannot exclude for instance that the antibodies

---

[377] This is not always true however, in that on the one hand the epidemiological results might be interpreted as incomplete data for what are otherwise exceptionless causal laws whereas, on the other, the results of many laboratory assessments are framed in probabilistic terms

[378] It is indeed what Williamson and Russo do in their (2007) paper. They analyse, among others, the Bradford Hill criteria for assessing evidence of medical causation and conclude that Williamson's account of *epistemic causality* – a neo-Humean account – does justice to what the medical claims amount to. No reference to the other, anti-Humean side of the debate appears in Williamson and Russo's paper.

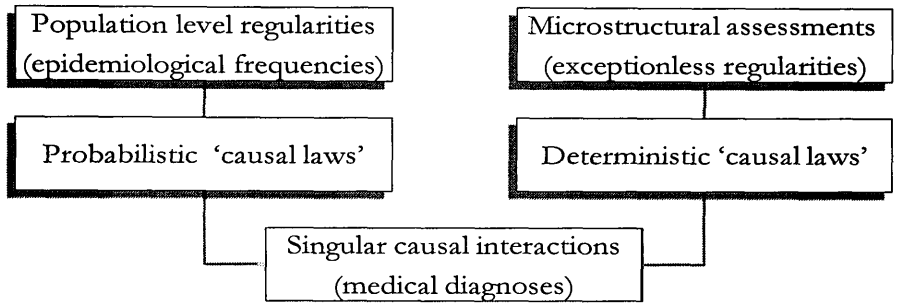
[379] Cartwright and Dupre(1988: 526-527) Cartwright (2007a: 58,65)

[380] Cartwright criticises in her (1999) the fact that huge financial resources are allotted to the breast cancer research investigating its genetic bases, while much less (financial) attention is paid to the research into the causal influence of the *endogenous* oestrogen levels on breast cancer, simply because 'the genetic programme is our best shot right now at a theory of everything' (*Cf.* Cartwright, 1999: 17).

[381] Cartwright (1999: 72)

that are generally held to block the neuro-muscular junction in myasthenia gravis actually improve the neuromuscular synapses and the muscular movements.[382]

FIGURE 21 CARTWRIGHT'S RATIONALE APPLIED IN MEDICINE



Contra Humean general causal claims – both the probabilistic 'laws' of epidemiology and the exceptionless 'causal laws' of physiopathology hold only for the context in which they were ascertained - these 'causal laws' do not have *universal* bearing [C 2007:58]

Contra Humean singular causal claims - neither the probabilistic 'causal laws' nor the deterministic assessment of laboratories can vindicate singular causal claims about interactions occurring in different causal contexts

Evidently, problems of ascertaining *singular* causation (i.e. the difficulties concerning medical diagnoses) should follow for the Humean from the problems involving the general causal claims. Recall Cartwright's simple example of a diet:

[382] In the case of the microstructural assessments we might have two sources of 'causal context' variation. One would be the different sets of causal factors showing up from one context to another, as different 'backgrounds' for the cause and effect under study – it is a sort of variation that these assessments share with the probabilistic, population level evidence. The other source would be the difference between the laboratory settings and the real life organisms. The medical laboratory conditions operate, of course, on grounds of a 'shielding' that eliminates (what an anti-Humean would call) a whole range of causal factors, interferences and stimuli. This latter sources of variation could evidently be invoked when questioning not only the Humean approach to medical 'causal laws' but also the move from 'causal laws' to singular causal claims. I shall return to this issue in §4.4

" [In] the more usual way...we insist that all admissible singular claims fall under a generic law...we take as a necessary condition for "this *C* caused this *E*" that it be generically true that *C*'s cause *E*'s... Now there *may* be some description, *D*, of me that fixes *in a law-like way* the efficacy of this diet for people who satisfy *D*. In that case we would have a reference class, picked out in a non question-begging way, in which *P*(weight loss/diet) > *P*(weight loss/no diet). But there may not be such a description. There may simply be individual variation. To insist that there is always such a description is to let a big – and insecure – metaphysical assumption guide our methodology. That I think is a wrong thing to do." (Cartwright 2005, italics original, underlines added)

The 'individual variation' that Cartwright alludes to is an extreme case of the difference in the causal context that her customary argumentation against the Humean strategy makes so much use of. The upshot of applying Cartwright's rationale to medical cases is that the 'causal laws' in medicine are not universal and one can doubt their use in vindicating singular causal claims. Given the presence in a particular organism of an etiological factor - probabilistically correlated (by whatever epidemiological means) to a certain set of symptoms, or identified (by whatever microstructural assessments) as part of a mechanism held responsible for the respective set of symptoms - one cannot justifiably infer that the factor produced that set of pathological properties instantiated by the particular organism in question.

In the next section, I shall present some of the problems that (dys)function based explanations have in dealing with the pathological realm. Discussing these problems is necessary because it will allow us to formulate two interrelated questions – what capacities are at work in medicine and what functions are at work in the *positive* functional explanations – which I shall try to answer by looking at what dispositions/capacities might be manifested in the functioning of items, in the case of diseased organisms.

§4.3 FUNCTIONS AND CAPACITIES

This section will revolve around an example of a particular pathological case. I shall use this example in §4.3.1 to outline the main problem that I take the dys-functional (or the failure-to-function) explanations to have, when it comes to singular organisms and their

symptoms - a problem that has to do with the distinction between positive and negative causes. Full-blown explanations, I shall argue, need to refer to positive causes, and function-based explanations can do this only if the *failure* part is dropped, as it were. That is to say, full explanations of pathological occurrences must cite the *functioning* (as opposed to the *dys*-functioning) of biological items.

Part of the aim of this section is to identify the sort of functioning that could be in place in pathological contexts. I shall then connect this line of enquiry with the issue of vindication of singular causal claims in medical circumstances (outlined in §4.2.3) and shall suggest in §4.3.2 that a plausible answer consists in viewing the functioning of items in diseases as the manifesting of anti-Humean capacities.

§4.3.1 *Problems of (dys-) functional explanations*

The example I am going to discuss is purposefully framed in the idiom of the case-studies that are frequently presented in the medical literature:

A 45 year old man is admitted in the emergency unit with complaints of shortness of breath on exertion, dizziness and palpitations. On observation, the patient exhibits rhythmic nodding of the head, alternating constriction and dilation of pupils, blanching and flushing of the forehead, quick filling and collapse of carotid arteries. The physical examination reveals tachycardia, wide pulse, left and down-ward deviation of the apexian shock and a descrescendo, early diastolic murmur in the axillary point. The paraclinic tests show on EKG ventricular hypertrophy, and systolic blood regurgitation in the left ventricle on cardiac ecography.

The blanching and flushing of the forehead, tachycardia, nodding of the head in time with the heartbeat, shortness of breath on exertion, dizziness, alternating constriction & dilatation of pupils etc., are all symptoms (and signs) of a valvular cardiac disease called aortic insufficiency (for simplicity, I shall henceforth simply call these symptoms and

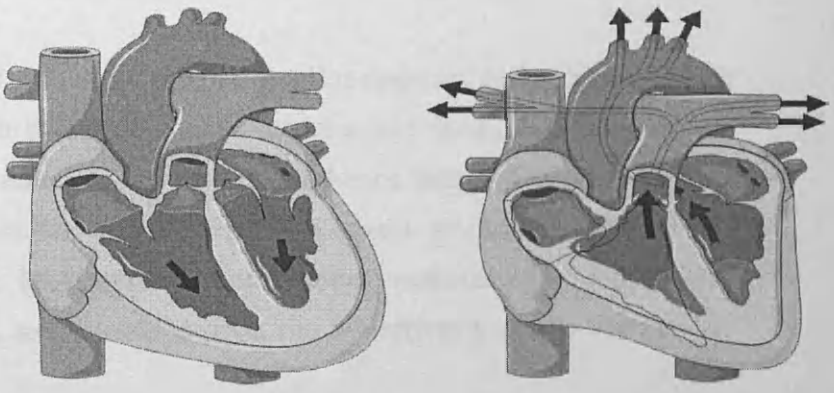
signs 'AoI symptoms').[383] The diagnosis of this disease is established using, *inter alia,* EKG and eco-cardiographic tests and its basic treatment consists in the medical administration of vasodilators and the surgical replacement of the aortic valve.

What function-based explanations could one provide for the symptoms the patient in question exhibits and how comprehensive could they be? Before addressing this question, I need to briefly present the standard medical textbook explanation of physio-pathology of aortic insufficiency (AoI).

Physiologically, the activity of the heart consists in alternating contractions and relaxations (systoles and diastoles), due to which the oxygenated blood coming from the pulmonary veins enters the systemic circulation and CO_2 loaded blood is redirected via the pulmonary arteries to the lungs. The circulation of blood through the heart in these specific directions depends crucially on a system of valves, which ensure that the oxygenated blood does not mix with the CO_2-loaded blood, that the contractions are used only for pumping blood in the arteries, etc.

Figure 22 **Ventricular diastole (left) and systole (right)**

As far as the systemic circulation is concerned, in a diastole the aortic valve (connecting the left ventricle with the aorta) is closed and the mitral valve (connecting the left atrium with the left ventricle) is open. The blood is then allowed to flow down from the left atrium to the left ventricle. In a systole the mitral valve closes down while the aortic valve is opened, and the blood is ejected into the aorta following the ventricular contraction.



[383] It should be stated that these are symptoms and signs of two conditions that are treated separately in the medical literature: aortic insufficiency and cardiac insufficiency. The former is taken as a valvular cardiac *disease*, as such, whereas the latter is discussed as a condition that emerges when almost all cardiac diseases progress towards their more critical stages. In other words, cardiac insufficiency can be taken as an advanced stage of aortic insufficiency (with the note that not only the aortic insufficiency undergoes this phase); see Arnold (2008)

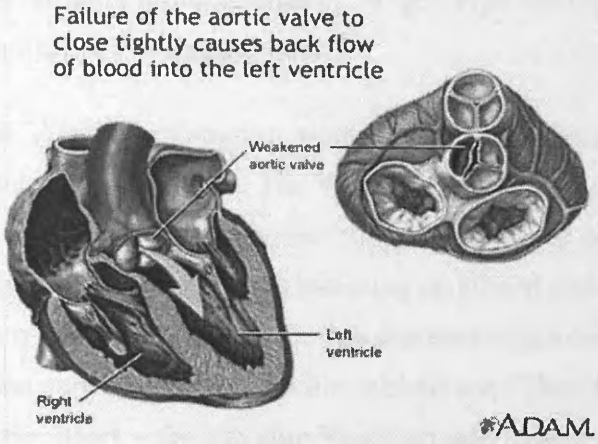
In a systole, the heart propels into the aorta 100-125ml oxygenated blood from the left ventricle (representing 60-65% of its volume) through the aortic valve. In the beginning of the diastole, when the left ventricle relaxes and the intra-ventricular pressure decreases, the aortic valve closes due to the high pressure in the aorta. Almost simultaneously, the mitral valve is opened and the blood coming from the lungs flows down into the ventricle. The filling of the left ventricle raises the pressure inside it. Consequently, at the end of the diastole the mitral valve is 'pushed' upwards and closed, while the aortic valve is opened when a new systolic contraction begins.

In aortic insufficiency (AoI henceforth), the aortic valve closes down incompletely at the end of the systole, such that some of the blood ejected in the systole leaks back into the left ventricle. In the diastole, the left ventricle thus receives both the regurgitated blood from the aorta and the blood flowing down from the left atrium. Hence, in the systole a greater volume of blood is pushed through the aortic valve, creating an increased systolic pressure in the aorta that decreases abruptly in the diastole along with the regurgitation. It is this difference between the systolic and diastolic pressure in the aorta that is responsible for the blanching and flushing of forehead, the wide pulse, the rhytmic nodding of the head in time with the heartbeat and the quick filling and collapse of the carotid arteries.[384]

The decreased 'net' outflow of blood (at the cerebral level) is responsible for dizziness and also for the activation of certain baro-receptors from the aortic notch due to which an increase in the tone of the sympathic nervous system and hence tachycardia occur. At the renal level, the decreased vascularisation activates a chain of hormones (rennin, angiotensin, aldosterone - the RAA system) that produce vascular constriction, the retention of Na and water and, to a certain degree, the hypertrophy of the ventricular walls.

[384] See Tanser (2007)

FIGURE 23 AORTIC LEAKAGE **(reproduced from Adam Health Encyclopaedia)**



Because the aortic valve closes down incompletely at the end of the systole, in the beginning of the diastole when the intra-ventricular pressure is low, a part of the blood ejected into the aorta flows back into the left ventricle. As a net result then, after a complete cardiac cycle, the blood from the left ventricle is ejected 'incompletely' into the aorta.

The decreased 'net' outflow of blood acts, *inter alia*, upon the sympathetic nervous system (producing tachycardia) and the RAA system (producing Na and water retention).

In the more advanced phases of the disease, (when AoI progresses towards the so-called cardiac insufficiency phase), the left ventricle becomes thicker and its volume increases, in order to cope with the increase in the volume of blood received in diastole; that is why we have a left and down-ward deviation of the apexian shock. The dilatation and hypertrophy of the left ventricle makes it relax less during the diastole and thus its diastolic pressure is heightened. When the diastolic pressure in the left ventricle reaches a certain level, it is 'transmitted' backwards to the left atrium, the pulmonary veins and further on to the pulmonary capillaries in which the air exchange is performed. That is how the shortness of breath occurs – the increased pressure in the pulmonary capillaries produces the extravagation of plasma proteins into the alveolar space, pulmonary compliance is then reduced and the pressure of O_2 in the exchange capillaries drops.[385]

The above represents what might be called the general physiopathological scheme of AoI,

[385] This happens in the advanced stages, in which *pulmonary oedema* occurs - one of the classic manifestations of cardiac insufficiency; see Arnold (2008).

described in very broad terms. Some further details will be adduced later. In this subsection I shall not discuss the status of the 'causal laws' that figure in this 'textbook' scheme but rather use it as a minimal backdrop to express what I see as the main problem of the *dys*-functional explanations.

What could explanations in terms of *failures* of functions tell us in this particular case? Evidently, depending on which account of function one adheres to, one could yield in this case either Cummins-type or Wright-type explanations. Let us look at the heart as one example of a biological item.[386]

The Wright explanation would be that the heart of this particular organism does *not* produce its W-effects. The W-effects, remember, are those which during a certain period in evolutionary history increased the chances of their organism surviving and due to which the heart's genetic basis was stabilized and maintained at the level of populations.. Even thought this heart in this diseased organism does do some pumping, the usual W-effects are manifested in a diminished way. Therefore, they do not qualify as the W effects of the heart - the genuine W effects are indeed the effects of the heart exercised upon blood circulation but only when they are *within* a certain range. Had the heart usually pushed blood with less than 50% ejection rate into the aorta and produced shortness of breath on exertion, for instance, it would have had much lower chances of being selected as a biological item in the *way* that it currently is.[387] Since the genuine W-effects are missing then in the organism of this person, his heart is *dys*-functional and this failure to function explains the disease and its symptoms.

The Cummins explanation would be based on the *typical* C-function - the function of the heart to produce the C-effects statistically determined in a population relevant for that individual organism. let the effects in question be called C1-typical effects, for reasons that will be clear very swiftly. These C_1-*typical* effects circumscribed by statistics,[388] construed as its contribution to the capacity of the circulatory system to distribute

[386] The following discussion could easily be applied to all the other items items implicated in the disease of the particular organism under scrutiny.

[387] See for instance Neander (1991), Nesse and Williams (1999) and Wakefield (2001). See also Sadler and Agich (1995) for a critical discussion, in the context of psychiatry, which focuses mostly on the disease/non-disease distinction and not on functional explanations, as such.

[388] e.g. cardiac output 70-75ml, ejection rate 60-65%, cardiac index 2.6 - 4.2 L/min*m^2, pressure in the ventricle in the mid-diastolic period 5-7 mmHg, etc.

oxygenated blood and eliminate wastes,[389] are not delivered in this patient. His heart simply *lacks* the typical C_1 function identified in the general population. We have then a failure of a C_1-function and pointing out this failure constitutes an explanation of his disease and symptoms.[390]

These two types of function-based explanations indicating the failure of the heart to deliver the W and C_1 typical effects, respectively, are somehow illuminating. Insofar as the notion of explanation is connected with *understanding*,[391] they do throw some light on the patient's state. Nonetheless, they do not constitute full-blooded explanations, in the sense mentioned in the previous section in which to genuinely explain a phenomenon is to pinpoint the causes whose effect it is.

In order to see this more clearly, let us recall some features of the basic explanatory role of functions, espoused in §4.1. In biology, the W-function of a token item T in a particular organism is taken to explain the selection of the item T, since tokens of the similar type-item produced the respective W-effects at a certain evolutionary point in the history of the lineage, which that particular organism is part of. On the other hand, the typical C_1-function of an item T belonging to a certain organism is taken to explain the capacity of a complex system (in a certain population) in which its typical C_1 effects play a role. Admittedly, these are in principle full-blooded explanations, since they explain certain aspects of organisms, by reference to their causes. The W or C_1 typical effects produced by items are taken in turn to act as causes producing the stabilisation in a certain population of the items' genetic bases (in the case of the W effects) and the employment of a certain complex capacity of a biological system (in the case of the C_1 typical effects).

When it comes to pathological situations though, what we see is that in the content of explanations, the place of functions is taken by *failures* to function (as explanatory

---

[389] I shall continue to employ Cummins's term of 'capacities' with regard to the behaviours exhibited by biological systems (explained by the Cummins functions) even though Cartwright uses the same term with respect to her anti-Humean powers. The particular sense in which this term is employed in the main text will always be clear from the context.

[390] See Juengst (1983: 131, 132). Juengst discusses the functional explanations that can be provided using Boorse's account of function, which, as I have mentioned in §4.1, could be plausibly classified as a type of Cummins account.

[391] Indeed, one of the ways that Hempel for instance tried to make sense of the notion of *explanation* was by connecting it with the psychological notion of *understanding*. See Hempel (1966)

elements for the pathological aspects - symptoms) and the place of the W effects and C_1 typical effects is taken by *the lack* of the W and C_1-typical effects, respectively. That is, the absence of certain causal factors is supposed to explain certain pathological aspects.

Now, it is true that in an extended account of causality, the absence of causal factors can be construed as a causal influence in its own way.[392] Nevertheless, such explanations founded on failures to function do not respond completely to the 'Why' question addressed to diseases, simply because the pathological elements of organisms do have positive causes on their own, beside the negative ones.

- Item T ---> W-effects ---->the maintenance of T
- Item T ----? lack of W-effects----? pathological aspects

- Item T ----> C_1-typical effect ---->the complex capacity of the system S in which T is part of
- Item T ----? lack of C_1-typical effects------? pathological aspects

I do not wish to suggest that the failure-to-function explanations are trivial. I suppose that a part of the contribution to understanding that these explanations seem to convey stems from the fact that, as discussed in §4.1, the C_1-typical and the W functions have been presented as capable of *identifying* diseases, via their failures. Whether they can do so is a long discussion. However, even admitting that the lack of the W or C_1-typical effects could help one identify and explain a certain state as *non*-healthy, I think that there is still a distance between *understanding* that an organism suffers from a disease, and specifically *explaining* the symptoms it exhibits.

In our 'case-study', we can perhaps understand that the person's organism is diseased. Yet, we do not have a straightforward answer as to why there is dizziness, shortness of breath, rhythmic head nodding, etc. Of course, conceding that the lack of a causal factor is still a type of causal contribution, one can point out that in that person's organism, a certain disequilibrium has shown up. After all, both the W and typical-C_1 effects of his heart are supposed to be part of a fine-tuned series of interconnections that are disrupted when these effects are not present. This is what function theorists like Neander are

[392] Even though there is no consensus on this matter; see Beebee (2001)

referring to, when adverting that in medicine the absence of the 'healthy' effects of items (the W-effects, according to Neander) explains what happens to organisms in pathological cases.

This sort of rationale does not get us very far from the general issue of identifying diseases though. Yes, a certain equilibrium is broken down, one could retort, but again, why does this organism exhibit blanching and flushing of forehead, alternating constriction and dilatation of pupils, etc. and not, say, wheezing, joint pains or Hippocratic fingers? In the end, what such explanations formulated in terms of failures of functions could tell us in this particular case is that, to put it informally 'The heart of this organism does not work well'. Is it sufficient?

The explanations revolving around failures of function (either W *dys*-functions or *losses* of C_1 typical function) constitute incomplete explanations of diseases. Again, I do not wish to suggest that they should be rejected. Nonetheless, it seems evident that a complementary project, would also give explanations in terms of *positive* causes. Only in this way one could explain precisely the emergence of symptoms, in what sense the survival of an organism is threatened, how the survival chances can be restored (by treatment), etc.[393]

In fact this is the direction of the 'pathological' explanations employed in somatic medical practice. The 'textbook' physiopathological rendition of AoI I have briefly framed at the beginning of this section appeals to *positive* causes.[394] This aspect is worth noting when looking for *function*-based explanations for the simple reason that the physiopathological, textbook-stance on AoI can in fact be viewed as a 'Cummins analysis' of the C-functions

[393] See Juengst (1983: 131, 132).
[394] It should be submitted that the terms 'dysfunction' and 'failure to function' are often met in the medical textbooks, in order to indicate the lack of certain effects. But the overall balance is overwhelmingly turned towards the positive causes side.

possessed by items in such a pathological circulatory system.[395]

I have pointed out in §4.1 that the appeal to the C_1 typical functions is an additional, 'external' element of the Cummins account of function, brought into discussion in order to respond to the Wright-ean challenge that an appropriate account of function should be able to *identify* diseases. These typical C_1 effects are supposed to be the 'healthy' ones, i.e. are supposed to be selected from a 'healthy' population (and parenthetically, huge problems attend this way of discriminating between health and disease).

However, it is by no means obligatorily for the 'pure' Cummins account of functions to refer to the typical C_1-effects in 'healthy' populations. The basic idea of this account is simply that, in order to attribute functions to biological items, one needs to locate complex biological systems and perform the so-called 'Cummins analyses'. To perform a 'Cummins analysis' is decompose the systems in question into elements, trace out the overall effects produced by each one of them and make sense of the contribution that the items make to the complex capacities exhibited by the overall system.

Whether these systems are pathological or not (on any other criteria), appears thus to be extraneous to the scope of a 'pure' Cummins account. Insofar as organisms with life-threatening conditions have systems possessing certain features of complexity, one can attribute C-effects (and hence C-functions) to the biological items participating in them, relative to the capacities exhibited by the systems under focus). Let us call these Cummins pathological effects the *typical* C_2 effects.

As I said, the C_2 effects in question, resulting from the Cummins *functioning* of biological

---

[395] There are in fact two interpretations of where the Cummins functions and Cummins effects are to be found. First, they can be taken to correspond, in general, to the analyses of the current non-evolutionary theory and/or second, to the study of mechanisms. See Craver (2001) for the latter interpretation and Amudson and Lauder (1994) for the former interpretation (in which, in fact the Cummins analyses are *also* said to describe the practice of evolutionary biology). The identification of a sort of Cummins functions and Cummins effects as being at play in the medical research spins off from the former interpretation – if what Cummins and his followers describe suits very well the biological practice, then it should also suit the medical practice. This proposal concerning the medical scope of Cummins functions, paradoxically, is often brought to the table by proponents of the Wright account (see for instance Godfrey-Smith 1993: 15-16) who are less preoccupied, however, with the explanatory project dedicated to medicine than they are with making the point that the Cummins functions could not be the *proper* functions. Since the Cummins functions can be identified in medicine, they argue, and the 'proper' functions should be able to 'identify' the pathological and make the distinction between disease and health, the Cummins functions could not be the searched-for *proper* functions.

items as well as the influence of these C_2 effects over systems (and ultimately, over entire organisms), could be taken to be the ones discovered by current medicine. That is to say, what general laboratory assessments, clinical trials, epidemiological studies, etc. are doing could in fact be viewed as amounting to 'Cummins analyses' which determine such C_2 effects (and C_2 functions) which can be employed to provide full blown function-based explanations of diseases.[396]

The C_2 effects of the heart appealed to in these explanations should be the hypertrophy and dilation of the left ventricle, the rising of the nervous sympathetic tone, the production of atrial natriuretic peptide etc. The C_2 effects of (certain types of cells, i.e. the juxtaglomerular cells, from) kidneys should be the synthesis of rennin and the subsequent production of angiotensin II. The C_2 effects of the aortic notch baro-receptors and the sympathetic nervous system should be tachycardia (and certain modifications in the *granula densa* cells). The C_2 effects of the suprarenal glands should be the secreting of aldosterone (and subsequently the Na and water retention and the hypertrophy and subendocardial fibrosis), etc. [397]

To be sure, in our imaginary 'case-study', this sort of Cummins explanations based on the C_2 *functioning* seems indeed to be in a better position than the *'failure* to function' approaches, when it comes to explaining the symptoms and pathological behaviours exhibited by the particular organism in question. These symptoms can be seen as resulting from the pathological capacities exhibited by its circulatory system, whereas the various C_2 effects of heart and the other implicated biological items could be taken as responsible for how the particular circulatory system at hand behaves.

- Item T → C_2 typical effects → pathological aspects (symptoms)

There are two important caveats that need to be stressed though. The Cummins analyses – which, I hypothesized, are to be identified in medical research, when general physiopathological schemes are circumscribed – could only have as results *typical* C_2 effects, taken as contributing in general to the capacities of circulatory systems of a

[396] See Geoffrey-Smith (2003)
[397] See Tanser (2007), Arnold (2008). More details will be provided in §4.4

certain type (say, AoI circulatory systems). On the other hand, the framework of the Cummins functions, even if much more liberal than that of the Wright functions, has certain restrictions that are not often mentioned by the function theorists, but matter to our discussion.

Correspondingly, we have two (related) worries, which can be roughly expressed as follows: how do we know that the heart and all the other implicated biological items produce in this *particular* organism the C_2 effects attributed to them by medical research and that these C_2 effects act as causes for symptoms? And, on the other hand, do these medical C_2 effects indeed constitute manifestations of *Cummins* functions?

In §4.3.2 I shall try to unfold these worries and provide some answers to them.

§4.3.2 *FUNCTIONING AS THE MANIFESTING OF CAPACITIES*

In principle, employing a sort of Cummins functions, related to what I have called the C_2 effects, should provide full-blooded explanations for the state of the organism of our case-study. This use of the Cummins functions, *prima facie,* does seem to be permitted by the general framework of this account. In diseased organisms we still have systems with a complex organisation and items that contribute to their behaviour. Since the activity of these systems is life-threatening and ultimately produces symptoms of diseases, the C_2 effects resulting from the 'Cummins analyses' performed in the medical research and the C_2 functioning seem strong candidates for being part of full blown functional explanations. Nonetheless, there are two (interconnected) worries about the very possibility of this sort of 'pathological' functions and their use in explanations for singular

organisms.[398]

The first worry goes back to the basics of the Cummins account of functions. For Cummins, as I have mentioned in §4.1, function ascriptions are supposed to be in fact attributions of *dispositions*, in the sense that the Cummins effects of items should be *manifestations* of dispositions possessed by the items in question.

Something may be capable of pumping even though it does not function as a pump (ever) and even though pumping is not its function. On the other hand, if something functions as a pump in a system **s** or if the function of something in a system **s** is to pump, then it must be capable of pumping in **s**. Thus, function-ascribing statements imply disposition statements; to attribute a function to something is, in part, to attribute a disposition to it. If the function of **x** in **s** to Φ, then **x** has a disposition to Φ in s. (Cummins, 1975: 757-8)

The dispositions of the items performing C-functions, however, are ultimately to be grounded in regularities. Cummins thus adheres to a Humean way of dealing with dispositions and causation. That is to say, in order to justify a (singular) causal claim of the sort - the item T of a such and such organism, in the presence of certain precipitating factors or stimuli, produced such and such Cummins effects (which in turn contributed to the their complex systems in a so and so way) - one should bring in regularities holding between items such as T, their C-effects and finally, the behaviours of systems. By going deeply enough, the regularities in question should turn out to belong, says Cummins, to the realm of 'pure physiology'.

'To attribute a disposition d to an object a is to assert that the behavior of a is subject to (exhibits or would exhibit) a *certain law-like regularity*: to say a has d is to say that a would manifest d (shatter, dissolve) were any of a certain range of events to occur (a is put in water, a is struck sharply). The regularity associated

[398] As I said, this use in medicine of the Cummins functions was suggested by several theorists, for instance, by Godfrey-Smith in his (1993: 15,16). Godfrey Smith (like many theorists discussing the Cummins functions) does not pay any attention to a crucial aspect of the Cummins account. To take another example, Craver, who devotes in his (2001) a considerable effort to defending Cummins's account against charges of triviality and showing precisely what type of systems (with what features of complexity) are related to the non-accidental Cummins effects, does not mention at all Cummins's dispositions; see Craver (2001). In the functions literature, Bigelow&Pargetter provide a very interesting *propensity* account of functions, which revolves around dispositions possessed by biological items; see Bigelow&Pargetter (1990: 332, 333). Their account however, as I have specified in §4.1, is a future-based variant of the *Wright* account. Interestingly, Bigelow and Pargetter are Humeans about causation and their (1990) is aimed at defending a Humean metaphysics. No specifics about causation are discussed however in the section dedicated to biological functions.

with a disposition - call it the dispositional regularity - is a regularity that is special to the behavior of a certain kind of object and obtains in virtue of some special fact(s) about that kind of object...by analyzing a disposition d of a into a number of other dispositions $d_1 \ldots d_n$, had by a or components of a such that programmed manifestation of the d_i results in or amounts to a manifestation of d [...] The biologically significant capacities of an entire organism are explained by analyzing the organism into a number of "systems" - the circulatory system, the digestive system, the nervous system, etc., - each of which has its characteristic capacities. These capacities are in turn analyzed into capacities of component organs and structures. Ideally, this strategy is pressed *until pure physiology takes over, i.e., until the analyzing capacities are amenable to the subsumption strategy* [i.e. until the analysing capacities are simple enough and can be reduced to regularities]' (Cummins, 1975: 758, 760-1, italics added)

Now, what would follow in the medical cases would be that, in order to ground the adverted C_2 effects (to be found in the physiopathological schemes), we should have some sort of dispositions, such that the C_2 functioning consisted in the manifesting of these dispositions. What could these C_2 dispositions be?

One immediate option is to stick to the terms of Cummins's original account, according to which (function-related) dispositions are connected with the level of pure physiology. Nevertheless, adopting this view has an untenable consequence. Given that the dispositions in question should be *reduced* to *regularities* of physiology, it would simply follow that the physiopathological C_2 effects were identical with the physiological ones – both types of effects being 'dissolved' in the same physiological regularities.

I have underlined several times in the present thesis that I shall not discuss the details of the distinction between health and disease (relying instead upon a broader distinction between those conditions that are life-threatening and those that are not). Yet, the conclusion that whatever causal interactions are in medicine, they are identical with the causal interactions from healthy organisms seems far-fetched,[399] even for my rather agnostic position on the (much discussed) precise distinction between health and disease.

Another option is to modify Cummins's original account and hypothesize that, even though we have 'pathological' dispositions present in the C_2 functioning, those

[399] To mention only the evident (contrastive) consequences, irrespective of the exact statistical life expectancy of a (human) organism with no aortic insufficiency, it is in any case higher than the expectancy of an organism with AoI reaching the stage IV NYHA of cardiac insufficiency (the latter being 1 year in 50 % of the cases; see Arnold 2008)

'pathological' dispositions are different from the physiological ones in that the former are to be 'dissolved' in the regularities of the diseases.

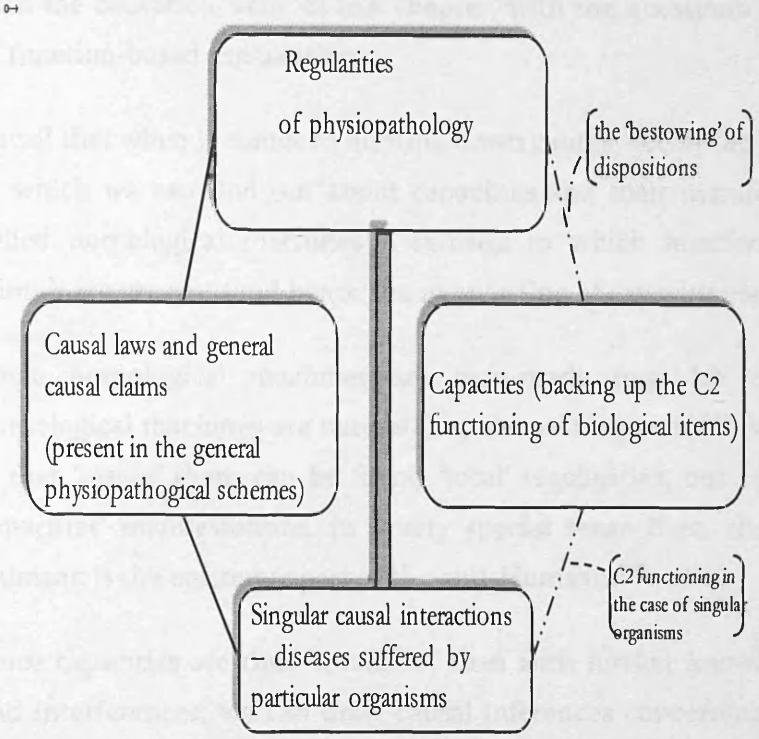
Yet, this alternative does not seem satisfactory either. On the other hand, what would follow would be a multiplication of dispositions – for each disease an item is engaged in, it would have to possess a different disposition. On the other hand, we should ask, what sort of regularities of physiopathology could be the ones in which the 'pathological' dispositions are to be dissolved? This brings us to the second worry.

The second worry is that in medicine we might not have the sort of regularities that would found general causal claims capable to vindicate singular causal interactions. To consider again our single case study, mentioning the C_2 pathological effects (to be found in the 'Cummins analysis' performed in the medical research) should amount to an explanation of the individual's state and the attending symptoms. That there is indeed a C_2 functioning of these items in this *singular* case is to be vindicated, in the Humean line that Cummins adopts, by mentioning dispositions and ultimately general regularities.

In other words, adopting the *explicit* idiom of causation,[400] the singular causal claim regarding the role of the C_2 effects in this particular organism should be vindicated by the 'causal laws' of physiopathology, where these causal laws are to be understood in a Humean vein, as derived from regularities. Informally put, we should know that the C_2 effects are there, in that singular organism, (produced by the respective items and producing in turn the manifestations from the level of systems) because we should have the 'causal laws' of the general physiopathological scheme of AoI.

[400]The two strategies that for Cummins are to be used in the 'Cummins functional analysis' - the subsumption strategy and the decomposition strategy (Cummins 1975) - were formulated in the 70's, using the neo-Humean idiom of dispositions, regularities, explanations, etc. which was supposed to make causation talk unnecessary. But arguably, what are at work in the case of the Cummins functions *simpliciter* are general causal claims, just like in the case of the C_2 functions are at work the general 'causal laws' of medicine. In the case of Cummins function attributions to particular organisms, there are involved singular causal claims (supposed to be justified, for the Humean, by the general causal claims). In the context of our discussion, the explanations citing these functions addressed to singular organisms are in fact singular causal claims (which are supposed to be vindicated by the general 'causal laws' of physiopathology).

FIGURE 24 THE HUMEAN STANCE ON THE C_2 FUNCTIONING AND THE ATTENDING CAUSAL CLAIMS



In the *explicit* idiom of causation, the singular causal claim regarding the role of the C_2 effects in a particular organism should be vindicated by the 'causal laws' of physiopathology, where these causal laws are to be understood in a Humean vein, as derived from regularities

An anti-Humean could make the case though that the 'causal laws' that underlie any general physiopathological scheme in somatic medicine are not universal and could not vindicate singular causal claims.[401] I have framed in §4.2 the rationale developed by Cartwright against Humean causes and the way this rationale might be taken to apply to the medical causes. In medicine, as everywhere else, one could adopt Cartwright's famous motto - no causes in, no causes out[402] – admitting that capacities should be reckoned with in our general and singular causal inferences.

[401] The rationale would be the same as the one I framed in 4.2..3 - it would cite the variation in the causal contexts given for instance different arrays of causal factors showing up from one population to another, the difference between laboratory setting and the real-life organisms, etc. The series of 'questions' that the anti-Humean could throw at the Humean strategy and 'causal laws' are also of the same type - does the decreased cardiac output produce the increased tonus of the sympathetic system and hence tachycardia? Well, we could have a population in which the sympathetic tonus is decreased by the lowered cardiac output. Does the increased pressure in pulmonary capillaries, given the rates of filtration, reabsorbtion and lymph flow, produce the extravasation of proteins in the alveoli? Well, we might have a singular organism in which the congestion of the pulmonary capillaries helps the alveoli remain intact.

[402] Cartwright (1989: 39)

I have asked thus in the previous section what capacities should be 'introduced' in our medical reasoning. I would like now to connect this question, which has specifically to do with the causation 'vein' of this chapter, with the questions surrounding the other 'vein' of function-based explanations.

Recall that when it comes to hunting down causes, according to Cartwright, the main way in which we can find out about capacities and their manifestations appeals to the so-called nomological machines – settings in which interferences are shielded, proper stimuli are present (and hence the manifesting of capacities is *regula*ted).

Some nomological machines are man-made (e.g. lab equipment),[403] while other nomological machines are natural (e.g. the solar system)[404]. What is common to them all is that 'inside' them can be found 'local' regularities, out of which we learn about the capacities' manifestations. In a very special sense then, the metaphysical path of the Humean is the epistemic part of the anti-Humean.[405]

Once capacities are thus 'localised', then with further knowledge of (abnormal) stimuli and interferences, we can draw causal inferences concerning singular interactions from the 'outside'. The question regarding what sort of capacities are at work in medicine should be then translated as: what medicine-related nomological machines do we have? Before attempting a response, let me sum up the two worries that I have framed in relation to the C_2 functional explanations.

The first one stems from the way Cummins construes the dispositions in the manifesting of which the Cummins functioning consists - these dispositions are said to be dissolved

---

[403] ....and their laborious construction requires singular causal 'information'- knowledge of stimuli, interferences, etc. which needs much experience but is in the end, says Cartwright (following Ascombe at this point) acquired because we can *perceive* natural necessity; see Cartwright (2002b)

[404] In the solar system, the interferences represented by all the other causal influences different from the gravitational one are naturally negligible; see Cartwright (1999).

[405] Whereas the Humean looks for regularities to reduce and ground *ontologically* the modal features of causation - recall capacities/powers are to be viewed as 'bestowed upon' by a certain type of constant conjunction - the anti-Humean looks for regularities in order to *know* about capacities. This is indeed a very special sense of 'path parallelism' – for the anti-Humean, the regularities from the 'inside' are local and not universal regularities. Moreover, when nomological machines are of the 'man-made' type, singular causal information about stimuli, interferences, etc. is essential in constructing them. In other words, that in a special sense the metaphysical path of the Humean is the epistemic one for the anti-Humean does not trivialise at all the Humean/anti-Humean distinction.

into the regularities of 'pure' physiology. Were the C_2 effects and the C_2 functioning dissolved into those pure physiology regularities however, it would follow that there was virtually no difference between what causally happens in ill and healthy organisms. That is, the causal interactions in physiology and physiopathology would turn out to be identical. On the other hand, were the C_2 effects taken to be dissolved into whatever regularities we might have in medicine, the dispositions each biological item has would be indefinitely multiplied, in accord with the potentially limitless number of diseases they could be implicated into.

The very attempt to dissolve the C_2 dispositions in question into the regularities of medicine is connected with the second worry, formulated from an anti-Humean standpoint - regularities cannot 'dissolve' dispositions and bear out the attending causal claims. In an explicit causal formulation, the singular causal claims regarding the C_2 functioning and the causal interactions at work between the C_2 effects, items and systems, in the case of a singular organism, cannot be vindicated using the general 'causal laws' of physiopathology, where these 'laws' are Humean, regularity based ones.

In line with the motto 'no causes in, no causes out', one can view the C_2 effects as manifestations of anti-Humean capacities.[406] Out of such capacities, 'inside' special contexts/settings, regularities spring, allowing us to learn about capacities' manifestations. With knowledge of capacities (and stimuli, interferences, etc.) one can formulate causal inferences about the singular interactions from 'outside'. But what could be the capacities at work in medicine?

My suggestion goes along the following lines: the capacities in medicine are physiological ones and the C_2 functioning results when these capacities are made manifest in

---

[406] I should emphasize, of course, that the second worry is formulated in the backdrop of Nancy Cartwright's criticism of the universality of the Humean 'causal laws'. Cartwright advocates the considering of capacities in our causal reasoning but her rationale is far from being a panacea: there are problems attending the anti-Humean program in general and her (particular) approach as well (see for instance Psillos 2008). I think though that at least the possibility of applying Cartwright's rationale in favour of capacities to medicine (in the context of the broader problems that functional explanations for diseases face) is worth investigating. As I stated from the very introduction of this chapter, my purpose is not to 'make a point' and adjudicate between the Humean and the anti-Humean options advanced in the literature, or between the various accounts of function on offer, but simply to connect the discussions over functional explanations in medicine with an area in the philosophy of science that has been neglected; see also §4.4.

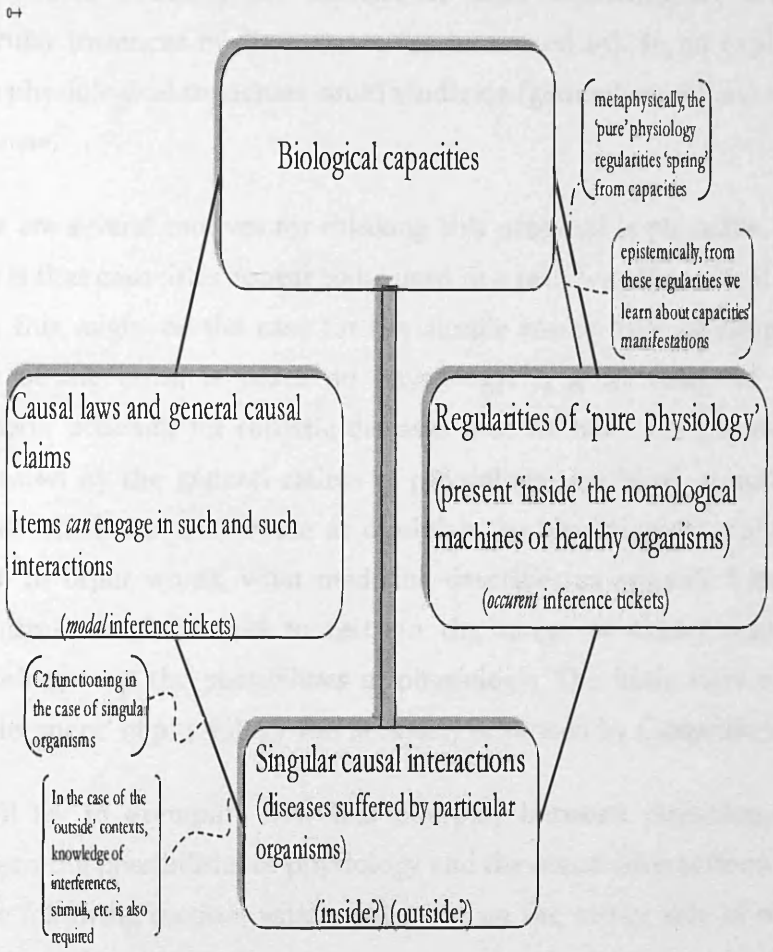
conditions where there are abnormal stimuli, interferences, etc. This requires 'healthy' organisms (in the minimal sense of organisms in which the survival is actively maintained) to correspond to 'nomological machines'. That is 'healthy' organisms should represent settings in which there is a perfect coordination between different internal components, such that the 'right' set of stimuli are present, interferences are screened off, etc., and the capacities of biological items are exercised optimally.

In healthy organisms, like in any nomological machine, regularities show up. A part of Cummins's original insight is thus preserved, because these are precisely the regularities that Cummins refers to when saying that his Humean dispositions are metaphysically dissolved into regularities at the level of 'pure' physiology. In the scheme I propose though, the capacities of biological items are just 'hunted down' at this level. In other words, the regularities in 'pure physiology' are just the epistemic platform for further causal inferences concerning different causal contexts.

Diseased organisms are ones in which items *preserve* their capacities, but the causal context changes in that the sheltering disappears and there are either interferences or exacerbating stimuli. In these circumstances, the capacities possessed by the biological items manifest waywardly – they are either exacerbated, diminished, unaltered or eliminated. Thus, the regularities from the nomological machine ('inside' healthy organisms) cannot *justifiably* be projected 'outside' (at the level of diseased organisms).[407] But the causal claims in singular cases can be vindicated, notwithstanding this change in causal context, by holding that capacities are still at work, and observing what sort of

[407] Recall that the crucial point of determining the capacities inside nomological machines is that regularities (resulting from their systematic manifestations) cannot be *justifiably* projected 'outside' nomological machines (and indeed this is the entire significance of Cartwright's insistence on *modal* inference tickets, which are different from the *occurent* ones). In the backdrop of our discussion, it seems evident that the regularities of physiology cannot be projected to the level of the pathological. This does not entail however that 'outside' there are no regularities - there *might* be, or there might *not* be (and Cartwright offers examples that cover both possibilities – think of the cathedral banknote situation *vs.* the example of charges in a specially framed magnetic field, both offered in her 1999: 27, 60). Indeed, in the medical realm there are certain diseases (especially the chronic ones) that exhibit a somewhat regular behaviour. In my scheme, this regular pathological behaviour occurs in an 'outside' context in which the abnormal stimuli and interferences, being present long enough and affecting the physiological capacities in a certain sense, produce regular manifestations. Thus, the causal laws of medicine should be taken to emerge out of such pathological regularities (construed as an epistemic and not metaphysical source). Insofar as causes *per se* are concerned, however, my primary interest is in singular causation in medical contexts. I shall not insist upon the general causal claims and their interpretation.

interferences, stimuli show up.



In a healthy organism the capacities each item has are perfectly correlated, in systems in which optimal stimuli show up and interferences are shielded

Diseases show up when this shielding is broken off, interferences appear, stimuli are modified, etc.. Thus, the fine correlations between different capacities are not present any more. In diseased organisms, we cannot *justifiably* project the regularities of physiology - the causal context is different. But we can use the *capacities* from biology.

FIGURE 25 **CAPACITIES IN MEDICINE**

This framework resolves the problems of the Cummins pathological explanations. On the one hand, I concur with Cummins, at work in the C_2 functioning of items, are dispositions, or capacities, intimately connected with the regularities of 'pure' physiology. However, unlike Cummins, I hold that these are anti-Humean capacities – capacities which depend on the regularities of 'pure' physiology on an epistemic and not metaphysical level and which are at work *both* in healthy *and* pathological conditions (when the regularities of physiology disappear). To cite the C_2 functions of items is to

provide full-blooded explanations of diseases, insofar as the various stimuli and interferences affecting the exercise of their capacities are also mentioned (for the particular instances of diseases we are interested in). In an explicit causal formulation, these physiological capacities could vindicate (general as well as) singular causal claims in medicine.

There are several motives for thinking this proposal is plausible. The most important of them is that capacities appear to be used in a tacit way in medical practice and research. I think this might be the case for the simple reason that physiopathology, in the broad sense of the term, is *based on* physiology. A great many of the physiopathological 'textbook' schemes for somatic diseases that we have are grounded on the possibilities delineated by the general claims of physiology, modified in accord with the abnormal stimuli which can exacerbate or diminish the physiological manifestations of biological items. In other words, what medicine describes as *actually* happening in the diseased organisms could be said to rest on the range of causal interactions delineated by physiology - on the *possibilities* of physiology. The basic view of medicine as a sort of 'development' of physiology was famously advanced by Canguilhem in his (1966). [408]

I shall try to exemplify how this interplay between physiology and physiopathology, between the possibilities of physiology and the actual interactions of pathology, is at work in the following section, which will insist on the causal side of my rationale concerning pathological explanations.

§4.4 CAPACITIES IN MEDICINE

The present section is oriented towards the causal vein of my argumentation on pathological explanations. I shall look here at an example of a biological structure that, as we shall see, is closely connected with our 'case-study' – the rennin-angiotensin-aldosterone system (RAAS henceforth), which has a particularly important behaviour in cardiac diseases. In §4.4.1, I shall use this example in order to describe the interplay between physiology and physiopathology I have mentioned in the previous section, in

[408] Even if Canguilhem was not concerned with the causation aspects I deal with in the present thesis.

connection with my hypothesis that organisms as nomological machines could do justice to the search for positive causes of diseases. Then, in §4.4.2, I shall return to our 'case study' and indicate how physiological capacities can be used to vindicate singular causal claims about pathological organisms.

§4.4.1 *WHAT ACTUALLY HAPPENS AND WHAT COULD HAPPEN*

The rennin-angiotensin-aldosterone system consists of specific enzymatic/hormonal pathways involving various organs (kidneys, the liver and the suprarenal glands) and could be viewed as part of the type of biological nomological machine I have hypothesized in the previous section - a biological setting which is disrupted in the cardiac diseases in a certain way, such that the causal roles of biological items are interfered with.

All cardiac diseases have as an ultimate stage a state called cardiac insufficiency and the RAAS plays a pivotal causal role in this pathological stage. The role in question, however, is arguably governed by the *physiology* of the RAAS. I shall describe in the following the main features of this physiology and then shall try to point out the links with cardiac insufficiency in general and aortic insufficiency in particular.

At the level of 'pure physiology', the RAAS acts upon the volume of circulating blood and the systemic vascular resistance in response to different environmental and internal variations, due to a series of hormones and enzymes - amongst which the most important are rennin, angiotension II and aldosterone.

Rennin is an enzyme produced by the so called juxtaglomerular granular cells, located in the wall of the afferent arterioles (the arterioles that bring blood for filtration in kidneys' glomerules). Its synthesis is stimulated by three factors: the decrease in the wall-tension of the afferent arterioles in question, the stimulation of sympathetic nerve endings located at glomerular level and the decrease in the concentration of NaCl in the tubular filtrate (the latter being sensed by another group of cells pertaining to kidney's nephrons

called *granula densa*).[409]

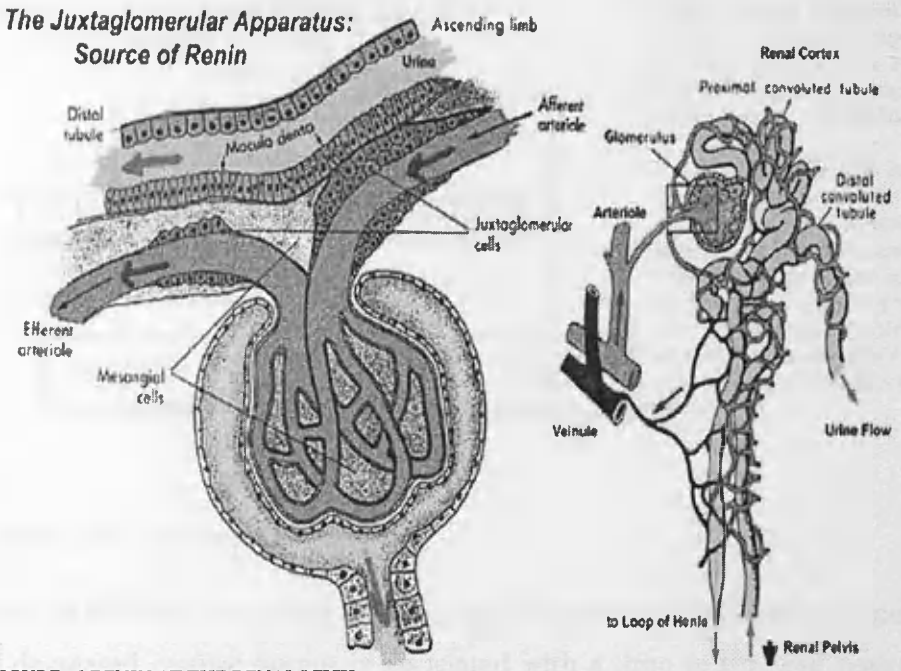
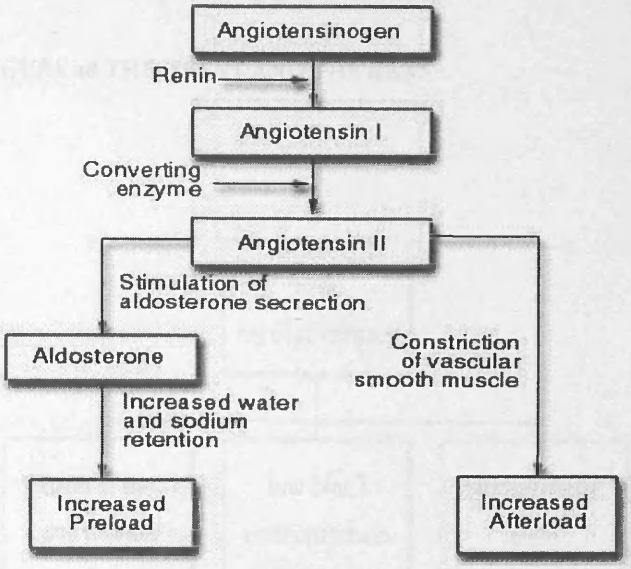


FIGURE 26 RENIN AT THE RENAL LEVEL

The juxtaglomerular cells that synthesize rennin are located in the wall of the afferent arterioles. These cells have sympathetic innervation and are also sensible to the variations of NaCl concentration in the distal tubuli, sensed by another group of cells called *macula densa.*

The result of renin's enzymatic activity is angiotensin II - a powerful vasoconstrictor that also stimulates the cardiac hypertrophy. Angiotensin II acts in turn on the adrenal cortex to release aldosterone, a hormone which at the renal level induces Na and water retention.

[409] See Lote (2000)



Renin breaks down a pre-hormonal chemical compound produced by the liver (angiotensinogen) enabling thus the release of angiotensin II, a powerful constrictor of vascular somatic muscles. Angiotensin II stimulates in turn the release by the suprarenal glands of aldosterone, whose main action is the increasing of Na and water reabsorbtion in kidneys' nephrons.

Angiotensin II has however many other important effects. It facilitates the norepinephrine release from the sympathetic nerve endings (thus contributing to the increase of the sympathetic tone), stimulates the release of vasopressin (a pituitary hormone which increases the fluid retention in the kidneys) and the thirst centres in the brain, etc.

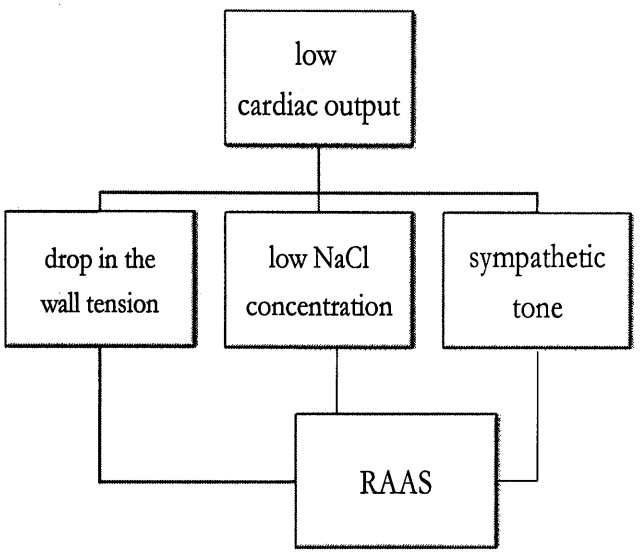
FIGURE 27 THE RENIN ENZYMATIC PATHWAY

One of the basic activators of the RAAS is the decrease of cardiac output. That is because a decreased cardiac output is associated with a drop in the wall tension of the afferent arterioles (due to the poor blood irrigation at the level of kidneys), a low NaCl concentration in the tubular filtrate (due to active reabsorbtion of NaCl whenever the renal irrigation is low) and an increased tone of the sympathetic system (due to the activated baroreceptors from the aortic notch).

The decrease in cardiac output is followed then - via rennin, angiotensin II and aldosterone's activity - by water and Na retention (which restores the circulating blood volume), vasoconstriction (which maintains blood pressure), further stimulation of the sympathetic system (which, *inter alia*, increases the heart' contractility), activation of the thirst centres in the brain (which complements aldosterone's effects), etc.

Informally put, when the blood delivered by the heart is lower in volume than usual (in various more or less transitory circumstances – dehydration, bradycardia, low Na intake, blood loss), the RAAS acts and restores the perfusion of tissues by contracting blood vessels, increasing the circulating blood/plasma volume and increasing heart's contractility. As soon as the cardiac output is restored, the RAAS is down-regulated.

FIGURE 28 THE HEART AND THE RAAS



At the level of 'pure physiology' the RAAS (the rennin-angiotensin-aldosterone system) acts *inter alia* upon the circulating blood volume and the systemic vascular resistance.

There are three main direct stimuli that initiate the enzymatic/hormonal pathways at the level of kidneys – a drop in the wall tension of the vessels that bring blood for filtration in the kidney's glomerules, a low NaCl concentration in the tubular liquid that results from filtration and an increased tone of the sympathetic nervous system in its renal ramifications.

When the blood delivered by the heart is lower than usual, the RAAS acts and restores the perfusion of tissues by increasing the systemic vascular resistance, heightening the re-absorbtion of Na and water in the distal tubuli and raising the heart's contractility. As soon as the cardiac output is restored, the RAAS is down-regulated

Now, in all heart diseases, when they reach the stage of cardiac insufficiency, the cardiac output decreases. It is not a transitory decrease however - the heart simply cannot deliver into the circulatory system sufficient blood. Take the case of aortic insufficiency - the disease that our 'case-study' is concerned with. As I have stated in §4.3.1, this disease results from a disruption of the aortic valve that allows for the ejected blood to leak back into the left ventricle. In the advanced stages, the aortic leaking produces a decrease in the cardiac output that cannot be compensated and the RAAS is accordingly stimulated. However, no matter what the RAAS does, the cardiac output (beyond the level of the aortic valve) is not restored and hence we have no down-regulation.

Put differently, as an ultimate result of the *'interference'* represented by the modified geometry of the aortic valve, the RAAS is *excessively* stimulated and this has unwanted consequences – the vessels are excessively constricted and we get hypertension (HTN), the blood volume is excessively increased and we get interstitial oedemas, etc. Relatedly, the heart undergoes hypertrophy, dilatation and subendocardial fibrosis. Thus, the intracardiac pressure is increased and then transmitted to the lungs. That is how patients with cardiac problems end up in emergency rooms with acute pulmonary oedema, to put it this way.[410]

Now, one could have an anti-Humean reading of this chain of pathological causal interactions, and by using the terms 'interference' and 'stimuli', I have already (more or less tacitly) adopted this reading. Importantly, the 'active' terminology I have used is in fact not far away at all from the one used in medical textbooks.[411]

But *explicitly*, what one could claim is that the pivotal causal interactions that take place in AoI cases are the ones delineated by physiology. We have first a range of causal interactions in which biological items *can* engage in. We have then interferences (in this case, the disruption of the aortic valve) that produce abnormal stimuli, which in turn exacerbate the manifestations of the biological capacities at hand.

[410] See Arnold (2008)
[411] See Tanser (2007), Arnold (2008), Sharpe and Swedgerg (1999)

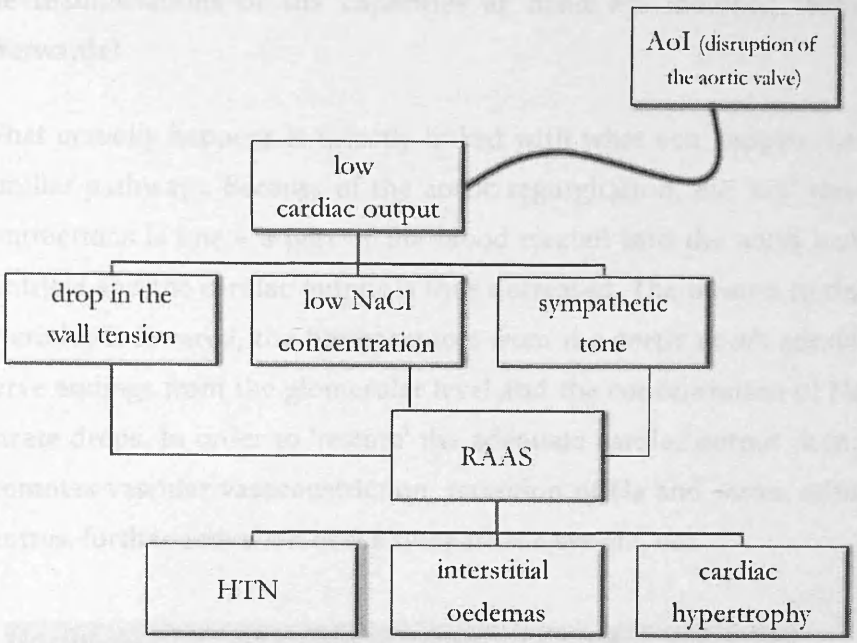


***FIGURE* 29 AOI AND THE RAAS**

At the level of 'pure physiology' the rennin-angiotensin-aldosterone system) upon the circulating blood volume and vascular resistance.

There are three main direct stimuli th enzymatic/hormonal pathways at the le – a drop in the wall tension of the vess blood for filtration in the kidney's glon NaCl concentration in the tubular liqui from filtration and an increased sympathetic nervous system in its renal

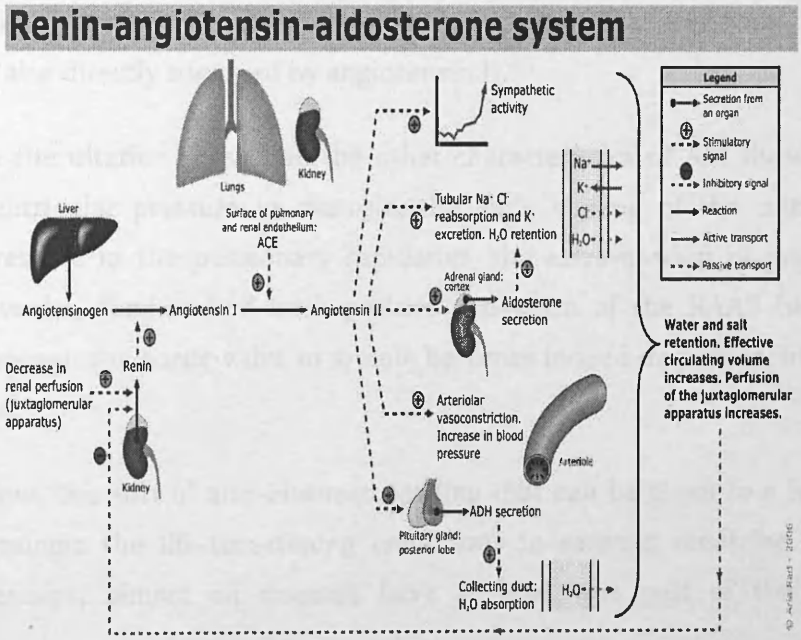
When the blood delivered by the heart i usual, the RAAS acts and restores the pe tissues by increasing the systemic vascul heightening the re-absorbtion of Na and distal tubuli and raising the heart's contr soon as the cardiac output is restored, th RAAS is down-regulated

The 'pure' physiology description of the interactions between the heart, rennin, angiotensin II and aldosterone tells us about the capacities that the items in question have. From the 'inside' of a nomological machine, in which the manifestations of their capacities are *regulate*d by the right stimuli and interferences are warded off, we learn a sum of modal behaviours, as it were. The heart *can* produce the lowering of pressure in the afferent arterioles, the increase in the sympathetic tone, the drop in the NaCl concentration at the renal tubular level, etc., rennin *can* produce angiotensin II, angiotensin II *can* produce vascular constriction, activate the thirst centres, induce the hypertrophy of the myocardial cells, etc., aldosterone *can* produce the retention of Na and water, etc.

Knowing what the heart and the RAAS *can* do permits one to make sense of the causal interactions that *actually* take place when interferences show up. When, in AoI cases, inside of the nomological machine constituted, *inter alia,* by the heart and the RAAS, a major interference shows up in that the aortic valve is disrupted, the stimuli triggering

the manifestations of the capacities at hand are modified. What *actually* happens afterwards?

What *actually* happens is directly linked with what *can* happen. Let me state again the familiar pathways. Because of the aortic regurgitation, the 'net' result of the ventricular contractions is low – a part of the blood ejected into the aorta leaks back into the left ventricle and the cardiac output is thus decreased. The tension in the wall of the afferent arterioles is lowered, the baroreceptors from the aortic notch stimulate the sympathetic nerve endings from the glomerular level and the concentration of NaCl from the tubular filtrate drops. In order to 'restore' the adequate cardiac output then, the activated RAAS promotes vascular vasoconstriction, retention of Na and water, stimulation of the thirst centres, further activation of the sympathetic system, etc.



What *actually* happens in AoI cases lies within the scope of (causal) possibility revealed by the nomological machine described in physiology. That the RAAS is actually stimulated in AoI cases is due to the exercise of one of the heart's capacities; that the Na and water are actually reabsorbed, blood vessels are actually constricted, etc. is due to the exercise of the capacities of the RAAS and its components.

***FIGURE* 30 WHAT ACTUALLY HAPPENS AND WHAT COULD HAPPEN**

To repeat, all these represent causal interactions that can coherently be viewed as spawning from the capacities of the 'pure physiology' description. It is just that the stimuli and thus the manifestations of these capacities are *exacerbated* by the interference specific for AoI. First, the stimuli received by rennin, angiotensin II and aldosterone are

excessive. That is because, *even though* the heart may deliver blood through the aortic valve at an optimum rate, the RAAS is *continuously* stimulated - informally put, the RAAS is triggered *as if* the left ventricle did not push enough blood through the aortic valve in systoles.[412] Had the interference not been present (i.e. had the aortic valve not been disrupted) the RAAS would have been 'informed' that the ejected blood through the aortic valve by the left ventricle, as such, was otherwise sufficient for the organism's homeostasis and (some of) its actions would have been down-regulated. Since the interference is present however, the RAAS acts in order to restore the blood circulating volume *and* heart's contractility.

Afterwards, the effects of (the overly exercised) capacities of RAAS act in turn as excessive triggers for the heart's own capacities. RAAS produces an increase in the vascular resistance, the so-called 'after-load' of the heart is accordingly heightened and this stimulates the myocardial capacity for hypertrophy and dilation. This myocardial capacity is also directly triggered by angiotensin II.[413]

In the ulterior phases, all the other characteristics of AoI show up in a row - the high ventricular pressure in diastole, the early closing of the mitral valve, the increased pressure in the pulmonary capillaries, the extravasation of proteins in the pulmonary alveoles, further feed-back positive activation of the RAAS (when the cardiac output through the aortic valve in systole becomes indeed decreased, in the final stages of AoI), etc.

Now, this sort of anti-Humean reading that can be given to a huge number of diseases, amongst the life-threatening conditions in somatic medicine. Indeed, except cancers perhaps, almost all diseases have at least one part of their mechanism involving

---

[412] In point of fact, this is a variant of AoI - the variant in which the displacement of the aortic valve is large enough to bring about quickly the cardiac insufficiency stage, in the context in which the volume of blood ejected through the aortic valve may (still) be unaffected. Of course, there are variants of AoI in which the displacement is smaller and thus the evolution of the disease takes longer. In such cases, by the time AoI reaches the stage of cardiac insufficiency (in which the role of the RAAS takes central stage) the volume of blood *as such* ejected through the aortic valve is also diminished. But in these variants as well, no matter the stage of the disease, the stimulation of the RAAS *worsens* the condition by bringing about hypertension for instance. That is why medicines such as ACIs (angiotensin convertase inhibitors) - drugs that block a crucial pathway of the RAAS - are administered in aortic insufficiency; see Sharpe and Swedgerg (1999)

[9] Angiotensin II induces among others the secretion of growth factors from the cardiac fibroblasts and endothelial cells. See Willenheimer *et. al* (1999)

diminishing and/or exacerbation of the causal relations from physiology. Graves' Disease is of course, another illustrative example.[414] Assuming that capacities are thus present in the medical realm, what is their involvement in medical diagnoses? That is, how could they justify causal claims about singular causal interactions taking place in particular organisms?

§4.4.2 *Capacities and singular causal claims for diseases*

I have argued that the physiopathological 'textbook' rendition of the role of the RAAS in aortic insufficiency is based on physiology, in the sense in which what *actually* happens in AoI (and in somatic diseases in general) could be said to rest on the range of causal interactions delineated by physiology, on the possibilities of physiology. Arguably, what we have in this interplay between the general causal claims of medicine (or its 'causal laws') and the general causal claims of physiology, are physiological capacities and certain 'abormal' interferences and stimuli that exacerbate or diminish the manifestations of these capacities.[415]

This is a sort of (anti-Humean) rationale that can ground *singular* causal claims

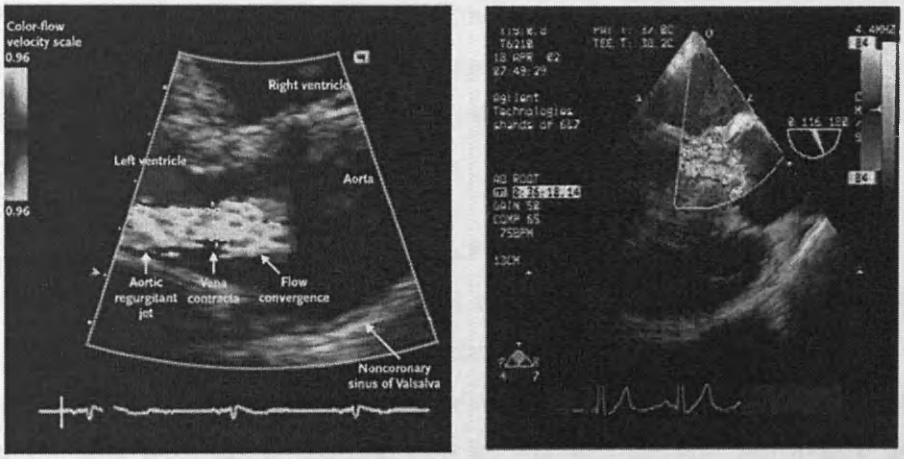
---

[414] Other salient examples that could readily be cited are peptic ulcer, Cushing syndrome, diabetes mellitus, glomerulonephritis, cardiomiopathy. As I have noted in the previous section, Canguilhem has actually provided a metaphysics of medicine which, even if not discussing explicitly the causation aspects I am concerned with in the present chapter, is grounded on the idea that the pathological emerges from an exacerbation or diminishing of the physiological; see Canguilhem (1991). As for the case of cancers mentioned above, it does appear that that they could not be readily assimilated to my scheme, simply because the morphological transformations they presuppose are often so dramatic that the link to the physiological background is lost (see for instance the case of Burkitt's lymphoma). Notably though, there are some cancers in which metaplasia simply presupposes that a type of tissue, in its pathological behaviour, assumes the physiological background of another type of tissue. For example, the small cell lung cancer induces high levels of cortisol in the blood by way of secreting a hormone called ACTH, which is physiologically secreted by the pituitary gland – see Chabner&Thomson (2008) and their discussion of paraneoplastic syndromes.

[415] That this anti-Humean reading can be viewed as delineating a sum of 'causal laws' for diseases is due to the fact that the 'context of diseases, even if situated 'outside' vis-à-vis the nomological machines of physiology, exhibit nevertheless a degree of regularity. When the nomological machines of physiology are disrupted and interferences, abnormal stimuli, etc. show up, the biological capacities do not manifest completely waywardly. In a sense, organisms with chronic illnesses themselves can be viewed as a *sort* of nomological machines that enable us to know about *stimuli* and *interferences*, and their influence on the manifesting of capacities. The full-blown nomological machines should however be taken to be the physiological ones, at least for reasons related to the principle of parsimony; see also the discussion of the worry concerning the multiplication of capacities, in §4.3.2

concerning organisms with the symptoms of AoI. Indeed, I shall try to show that it is a rationale that is also tacitly present in the somatic medical practice and research when it comes to medical diagnoses, i.e. when it comes to causal conclusions about particular instances of diseases.

Let us make a bit more vivid the hypothetical scenario around which a part of this chapter has revolved. Envisage the patient of our 'case-study' with the sheer symptoms and signs of AoI (shortness of breath on exertion, dizziness, rhythmic nodding of head, quick filling and collapse of carotid arteries, etc.) entering into an emergency yard. The medics perform the most important paraclinical test, the eco-cardiography, which shows *inter alia* the leaking of blood from the aorta.



/Transesophageal ecographies of AoI hearts, catching the aortic regurgitant jet, the hypertrophy and dilatation of ventricular walls, the high ventricular diastolic pressure, etc.

FIGURE 31 CARDIAC ECOGRAPHY (REPRODUCED FROM JANELLE, 2003)

Usually, this is enough to put a diagnosis. Of course, if the patient is really in a critical state, further tests are performed - the pressure in the pulmonary capillaries is ascertained using an intra-atrial catheter, an EKG is used to check (among other things) the ventricular hypertrophy, a urinary test assesses (among other things) natriuria, etc. The question we need to consider is – would all these tests be enough for a Humean? Would they be enough for the Humean to infer that, in this particular instance, the modified geometry of the aortic valve has enabled the chain of interactions whose outcome are the symptoms and of his overall condition?

The Cartwright-inspired critique that I have presented in §4.2.2 and §4.2.3 would be that,

from a Humean point of view, all these tests show *nothing*. Insofar as a Humean is concerned, a proponent of Cartwright would say, we might have there a case of accidental succession of one factor (the putative effect or set of effects) following spatio-temporally another factor (the putative cause or set of causes), irrespective of whether we discuss here about the heart or any other implicated item. These tests assess a case involving singular interactions and the medics draw conclusions about the particular organism at hand. A Humean needs however (universal) regularities.

We have the general 'causal laws' of the physiopathological schemes, which, in practice, guide the medics in inferring what *actually* takes place in that singular organism and which are based on endless population level assessments and microstructural laboratory research. But ultimately, if no regularities of the sort demanded by the Humean can be found or justified, then the medics' conclusion rests in the end on more or less tacit anti-Humean grounds. And indeed, one could argue that the regularities demanded by the Humean are missing, in that in the case of this patient we may have a *different* causal context.

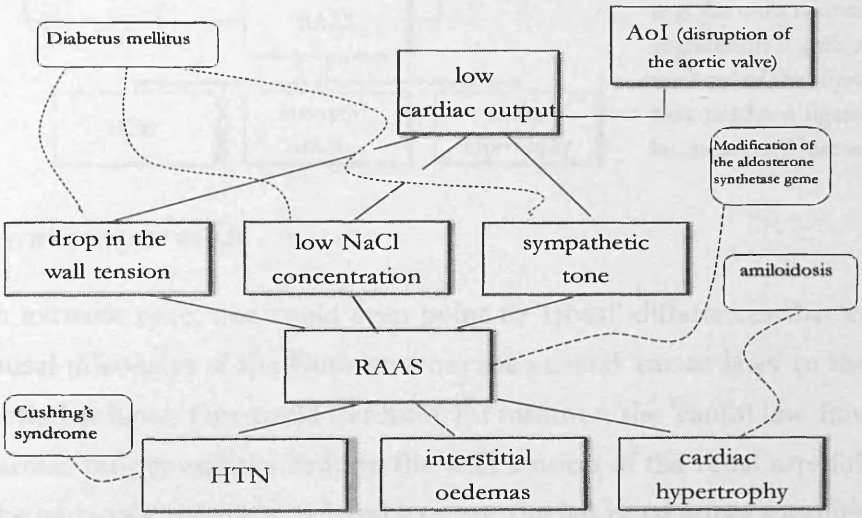
For one thing, there might be in place a combination of causal factors and stimuli that does not resemble any of the causal contexts in which the general physiopathological scheme was established.[416] Suppose that the patient of our 'case-study' suffers not only from aortic insufficiency (AoI) but also from diabetes mellitus II, Cushing's syndrome, amiloidosis and has a modification of the aldosterone synthetase gene enabling an excessive aldosterone synthesis (familial primary hyperaldosteronism form 1).

The factors that are present in these conditions can also act on the same sites that the factors of AoI are said to act – diabetes mellitus influences the sympathetic tome and the NaCl concentration in the tubular filtrate, amiloidosis produces cardiac hypertrophy, Cushing's syndrome produces hypertension, the faamilial primary hyperaldosteronism has as consequence, of course, increased aldosterone synthesis. [417] The precise combination of the causal factors present in these conditions might not have been

[416] Of course, what I try to frame above is what could count, for this imaginary patient, as a variation in the causal context that would mark the difference between the 'internal validity' and the 'external validity' of our causal claims discussed by Cartwright; see Cartwright (2007:61) and also §4.2.2

[417] See Kasper *et. al* (2008)

studied in any sort of laboratory research or/and population level assessments. If the background constituted by this array of factors was not explicitly present though, then any imperfect regularities established (by whatever means) at a population level and the attending Humean probabilistic general causal claims - just as any exceptionless regularities (established by whatever means) at a microstructural level and the attending Humean deterministic causal claims - cannot rule out the possibility that the disruption of the aortic valve was not actually a preventative for the symptoms that subsequently emerged.

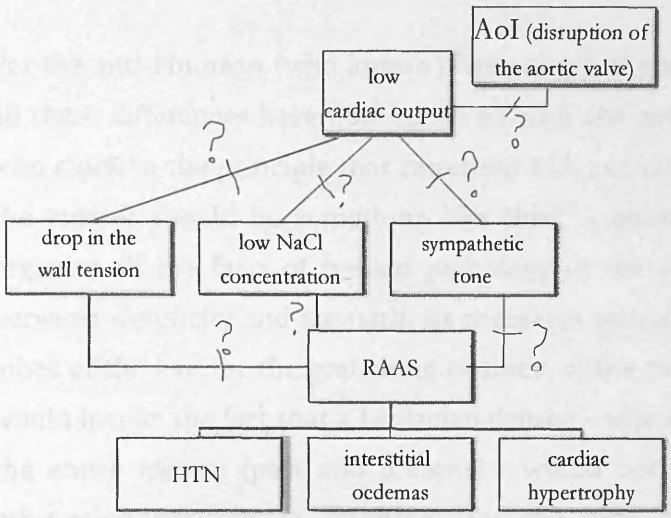


A Humean could not rule out that the disruption of the aortic valve was actually a preventative for the emergence of the symptoms

FIGURE 32 **A CAUSAL CONTEXT CONSTITUTED OF A BACKGROUND WITH A DIFFERENT ARRAY OF CAUSAL FACTORS.**

One does not need to resort though to such generous combinations of diseases (and causal factors) to argue that this patient may not behave as others have. One could point out that the causal context of this patient is unique only by mentioning, on the one hand, that in the population level studies the participants are randomised. On the other hand, one could mention that there are differences between any more or less shielded laboratory settings researching the microstructural aspects and real-life human organisms - in the 'sheltered' conditions of laboratories, causal influences that show up in real-life

situations are eliminated. [418]



Each of the 'causal laws' of the AoI scheme could be questioned with regard to its scope over individual organisms, even by appealing to 'trivial' differences constituting a 'different' causal context.

The interactions inside the RAAS (e.g. the ones connection angiotensin II with aldostrone, at the level of the suprarenal glands) have not been figured here but can be questioned just as well.

FIGURE 33 WHAT CAUSES WHAT?

On an extreme note, one could even point to 'trivial' differences that could undermine the causal inferences of the Humean from the general 'causal laws' to the singular causal interaction at hand. One could question, for instance, the 'causal law' having as relata the low cardiac output and the drop in the wall tension of the renal arterioles by indicating that the particular organism at hand has, say, the left renal artery spinning off the aorta in the thoracic region. [419] To take another example, one could question that, in this organism, the RAAS produces cardiac hypertrophy by pointing out that the liver (the organ producing angiotensinogen - the peptide precursor of angiotensin II) has two lobes instead of three, [420] etc.

Of course, these are trivial differences, even for a lay eye, not to mention a medic or biologist. Nonetheless, we should ask, from what point of view are they trivial? Why do not the number of lobes of the liver or the branching position of the renal artery (when it

[418] Examples here could be the microbiological studies *simpliciter* that isolate systems, organs, tissues, cells – down to the level of cellular components. Animal experimentation could also be adduced; see Cartwright's discussion of laboratory conditions and Galilean idealisation in her (1989: 185-188, 227, 228). Randomisation in medical trials is discussed in Cartwright (2007a:58-68).

[419] This is a variant described in Fernet *et. al* (1987)

[420] The third being the so called Riedel's lobe - an inferior projection of the right lobe on the right of the gallbladder; *Cf.* Gillard (1998)

comes to producing of rennin or the emergence of cardiac hypertrophy) count as interferences, or as having a causal influence of whatever sort?

For the anti-Humean (who knows physiology!), the answer is (relatively) simple: because all these differences have nothing to do with the exercise of capacities.[421] For a Humean who stuck to the principle that causation talk can ultimately be grounded in regularities, the answer should be something like this[422] - because in the Best System that would organise all the facts of human pathology in the entire history with the best balance between simplicity and strength, its theorems would/should not mention the number of lobes of the liver or the branching position of the renal artery. Alternatively, the answer would invoke the fact that a Laplacian demon - who knew everything that occurred along the entire history (past and present) - would not count any liver lobes or any aorta bifurcating position in inferring (for the case of an AoI organism reaching the insufficiency stage), from the decrease of cardiac output, all the consequences mentioned in the general physiopathological scheme of AoI.

In their decisions, the medics however do not consult any Laplacian demon or appeal to any best systems organising the facts of the entire history in a such and such a way. This is, if you like, a *methodological* point and not a *metaphysical* one.[423] It might well be the case that, in the end, the Humean metaphysical position is just as coherent as the anti-Humean one -'the spoils to the victor' as David Lewis used to say[424] - and a further discussion of this dispute is beyond the scope of this chapter. But when looking at the actual practice of medicine, the distinction made by medics between what is trivial and what is not helps one in interpreting the conclusions drawn in medical research and diagnoses as tacitly involving an anti-Humean stance.

It has often been argued that physicists act in their experiments ignoring the Humean

---

[421] On this Humean reading, knowing physiology is knowing about the physiological capacities and what stimulates (or interferes) with their manifestations. This knowledge is translated into the medical research and practice. Arguably, medics possess a knowledge of stimuli and interferences that is not limited to the textbook information but is also accumulated in practice and can be used to differentiate between particular cases and particular patients. The latter experience is important because it cannot be codified in general rules of thumb. It is not a 'catcher', to use one of Cartwright's suggestive expressions (Cartwright 2007[a]: 62-64), which could provide us deductively the causes at work in individual cases.

[422] See for instance Papineau (1991: 409), Russo and Williamson (2007: 13) Schrenk (2008)

[423] See Carwright's reply to Lipton's comments on 'The Dappled World' in Cartwright (2000[c]: 271-279)

[424] Lewis (1995)

worries on causation.[425] What I wanted to suggest in the last part of this chapter was that the same is the case in medicine. Any treatise of physiopathology dwells on the causal interactions of a physiology treatise - the huge majority of the *actual*-s described in physiopathology rest on the *can*-s of physiology and knowledge of interferences and stimuli. In the same vein, physicians and medical researchers, when moving from the level of laboratories and population level assessments to the one of particular organisms, work using anti-Humean assumptions. They assume a range of causal interactions which items *can* engage in, and adjudicate which of the potential effects (and to what degree, how intensely) shows up in singular cases, using contextual knowledge of the particular stimuli present in each case.[426]

§4.5 Concluding remarks on functional explanations

I have outlined here an approach to functional explanations that is in accord with the account of natural kinds of diseases developed in chapters 1-3. Let me recall very briefly the rationale that represented the starting point for the present chapter. Since kinds of diseases are circumscribed by determining properties (biological properties attended by causal powers) and to explain is to pin down causes, it seems to follow that the explanations for diseased organisms should be searched for at the level of the determining properties characterising their kind. That is to say, the causal powers attending these properties should explain the interactions of biological items and the emerging manifestations (symptoms) exhibited by the organisms under focuses. On the

[425] For instance Lowe (1989) and Armstrong (1997) have drawn such a conclusion with regard to the crucial (non-repeatable) experiments in physics. Cartwright has made a similar point when discussing the Stanford gravitational experiment in her (1989) and (1999).

[426] In the philosophical literature there are no systematic treatments of the role of anti-Humean powers in medical causation. In fact, there are very few papers that treat specifically medical causes and mention the Humean/anti-Humean distinction (excepting here the frequent, dilettante attempts of epidemiologists to refer to natural necessity; see for instance Lipton and Ødegaard (2005). Moreover, most if not all of them adopt a Humean view of medical interactions (e.g. Whitbeck, 1977, Marcum, 2008). Very recently for instance, Williamson and Russo have applied in their (2007) a neo-Humean account (i.e. the epistemic account) to medicine, concluding that 'an epistemic account of causality is *required* to capture the full complexity of [medical] causal evidence' Russo and Williamson (2007:13, italics added). No mention of the other, anti-Humean side of the debate was made in their paper though. In this context, framing an anti-Humean response or alternative - even if in a sketchy manner, as I did - might represent a worthy enterprise.

other hand, functions seemingly represent an inextricable part of biology and medicine, which any philosophical approach to explanations needs to reckon with. How can one deal with the functional explanations of diseases then?

I have tried to provide an answer, on the one hand by manoeuvring the pair of distinctions function/function failure and positive/negative causes, and on the other hand by interpreting the functioning of biological items as the manifesting of capacities. In the following, I wish to add a few remarks about some particular parts of my argumentation, and also a conjecture.

In the final part of this chapter, a specific assumption I have adopted was that in the realm of the pathological, the Humean cannot find the appropriate sort of regularities in order to justify the general causal claims advanced in medicine. The respective claims, I have shown, could be vindicated in the framework of biological capacities - capacities revealed at the level of 'pure physiology' and in the exercise of which the functioning of items consists. The level of 'pure physiology' was taken to be one at which the exercise of these capacities (and the functioning of items) is regulated by the right stimuli and all the interferences are warded off.

Now, claiming that, by the rigorous Humean standards, the sort of regularities that are found in medicine cannot ground general (as well as singular) causal claims, evidently, does not entail that the medical domain is a realm of chaos, in which no regularities whatsoever may be found (using this time 'regularities' in a broad sense).[427] Had diseases not exhibited some sort of patterns, had not symptoms grouped themselves in syndromes, etc. medical science would not have been even conceivable. Medicine has its sort of regularity and stability, and has a plethora of causal claims that are successfully used in treatments and further research. It is simply that, one could argue, the stability found in medicine is not sufficient to ground a Humean interpretation of these causal claims. There is always a possible variation in the causal context. The fact that in spite of this variation, the 'causal laws' found in the physiopathological schemes work and are fruitfully used in research and treatments simply underlines a point I have tried to make clear in § 4.4.2 – tacitly, medics employ anti-Humean assumptions.

[427] Recall that, according to Cartwright, not even physics and chemistry exhibit the sort of regularities that would fit the Humean stance on causation; see §4.2.1 and Cartwright (1999).

SOME FINAL REMARKS

I have mentioned in chapter 1 that natural kinds are the modern heirs of what the scholastic philosophy, by taking up a mainly Aristotelian (but also Platonic) scheme, construed as Ideas in God's mind. From this point of view, that diseases could be natural kinds appears almost as a blasphemy. When creating the world God could not have considered that diseases need to have a place among the substances constituting the reality, as it were. Indeed, it was the Aristotelian concept of substance that was involved in scholastics' endeavour to reconcile the Greek ideal of an eternal universe with God's omnipotence by introducing the paradigmatic Ideas as intermediating the Creation. Diseases are imperfections, distortions of (substantive and non-substantive) universals' instantiation in the sublunar realm, a medieval would have said.[428] In Locke's (already anachronistic) terms, one could say that diseases are dependent entities, modes, and implicitly suggesting that they qualify among the substances is a 'category mistake'.

On this side of the discussion, a proper answer from the defender of diseases as natural kinds would have to consider the history of philosophy and the way in which, from the seventeenth century onwards, the concept of substance has begun to lose its original significance.[429] I could not have tried to tackle such problems in the present thesis, which was strictly concerned with an analysis of the medical realm vis-à-vis natural kinds and functional explanations, with the conceptual means and strategies of the contemporary philosophy of science. In my argument that diseases are natural kinds I tried, however, to consider as much as I could this 'blasphemous' category mistake. I did this by offering a lot of space for discussion to the intuition of carving nature (and its current expressions and possible substantiations) and also to the phase/non-non phase distinctions, in the multiple ways in which it seems to threaten the placing of diseased organisms into the category of natural kinds of objects.

In fact, my argument in chapter 2 that certain natural kind requirements are not justified could have been more simply expressed by merely emphasizing that criterion IV plays a crucial role in the evaluation of several other criteria. I chose instead to integrate the

[428] See Nandler (1998) Klima (2008), Funkenstein (1989)
[429] See Lagerlund (2009)

argumentative vein concerning the diachronic identity of kind members within the interrogation over the intuitive burden that attends contemporary discourses about natural kinds. I then offered a divide as to what these intuitions could possibly mean (if we distance ourselves from the initial Aristotelian/Platonic/medieval context and look at present-day metaphysics of science). Indeed, my point could be reformulated by saying that, if one would want to prove that there we have a category mistake in viewing diseases as natural kinds (without appealing to God's ideas or whatnot), one would have to bite the bullet and assume explicitly the identity sense of carving nature at its joints. Similarly, if one would want to show that diseases are fundamentally processes and the 'static' framework of natural kinds of objects is entirely out of place, one would have to find entirely new ways of drawing the phase/non-phase distinction.

With this identity caveat, I have shown that between (life-threatening) conditions in somatic medicine and the exact science kinds we do not have an ontological gap but a difference of degree. Further on, I have shown that we could provide adequate *functional* for diseased organisms construed as natural kind members, by way of appealing to Cartwright's notion of capacity. It should be said that Cartwright situates her capacities as modern heirs of Aristotelian natures.[430] Given that, of course, the very notion of function is Aristotelian, I could say that one additional reason why the enquiry in the first and second part of my thesis is unitary is that it amounts in a sense to a (tentative) re-conceptualising of traditional Aristotelian categories and giving them new uses.

Leaving the history of philosophy aside, Cartwright is also an important author for the first part of my thesis because my claim that no ontological gap separates disease kinds and exact science kinds is in the trail of the 'dappled' picture of the world that Cartwright argues in favour of, in her attempt to reject the hierarchical view of sciences.[431] While being very sympathetic towards and actually admiring her views, however, I have not adopted (and did not need to adopt) in my discussion of natural kinds her starting points and argumentative aims. Cartwright mainly discusses causation, seems to adopt an 'emergentist' view of special science properties and is a staunch anti-Humean.[432] On my

---

[430] See Cartwright (1989: 187-200, 1999: 77-104)
[431] See Cartwright (1999)
[432] See her reply to Peter Menzies's suggestions, in the Philosophical Book's issue dedicated to her work; Cartwright (2002c) and Lipton (2002)

part I have only assumed supervenience for medical properties and the anti-Humean stance was cautiously adopted, as an conceptual hypothesis that could add to the plausibility of functional explanations for disease kind members. This cautiousness has its roots, of course, in the limited scope of this thesis, but not only.

To appeal to a rather exotic illustrative treatise in the present context, in the Introduction to his 'Liberalism',[433] John Gray, a defender of this political doctrine, stated that he wanted to offer an exposition of liberalism that could be useful both for its friends and enemies. Gray's respective depiction of both the strengths and weaknesses of a position is an impressive example of fairness in argumentation, which I have tried to follow, on a more modest level. Thus, someone who wanted to preserve the classical (medieval) connotation of substance could actually use the first chapters of this thesis in order to deny the existence of natural kinds in medicine, simply by following the identity caveat I have formulated. [434] The epistemological difficulties I have portrayed to the identity sense of carving nature seem to me extremely serious but the underlying metaphysics is coherent and of course, I could not exclude that these tremendously intricate difficulties could be overcome. This does not mean that my conclusion is weakened. If these difficulties are not overcome, again, diseases are natural kinds.

The same is the case for my application of Cartwright's anti-Humean scheme in order to make sense of functional explanations for disease kinds members. Cartwright often uses arguments that have to do with the epistemology and practical methodology of science in order to advance metaphysical claims about the nature of causality.[435] However, I have mentioned from the beginning of chapter 4 that my purpose is not (and could not possibly be) to adjudicate between Humeanism and anti-Humeanism in general, and even less to point out what nature the causality governing the medical realm has. What I have tried to do was to show that *if* we're trying to make sense of functional explanations for diseased organisms construed as members of natural kinds, then Cartwright's scheme is a suitable working hypothesis. This hypothesis suits the special sciences, is coherent

---

[433] Gray (1995)

[434] In point of fact, my caveat regarding the identity of kind members is closely connected with Aristotle's view of the relation between primary and secondary substances See Gobbo (2000:84), Dahl (1997:233), Ayers (1991:20)

[435] For criticism, see for instance Psillos (2008), Papineau (1991).

with practical judgment of medics the medical practice and offers a background our intuitions that 'something has gone wrong' (and is exacerbated and/or diminished) in diseased organisms.[436]

However, any argument as to what causality in medicine *is*, would have to take into account not (only) the methodology of medicine and our conceptual hypotheses on various levels (as for instance in relation to the natural kind level). Such an argument would also have to consider the intricate metaphysical dilemmas that concern the nature of properties, their association with causal powers, the identity of these causal powers, etc.[437] In this sense, again, the general description of the metaphysical positions involved in the Humean/anti-Humean dispute from §4.2 as well as the clarification of Cartwright's position from § 4.3, could be used by one in order to argue that actually, Humeanism *only* suits the medical science. But again, such an argument would have to tackle the intricate metaphysical dilemmas I have mentioned. Otherwise, anti-Humeanism remains a powerful hypothesis for causality in medicine.

Word count: 98, 639 (bibliography included)

[436] Given that, as I have stated in the end of § 4.4.2, in the philosophical literature there are no systematic treatments of the role of anti-Humean powers in medical causation, I thought that looking at how Carwright's views suit medicine is also worthy enterprise.

[437] See Williamson and Russo (2007) who discuss the causes in medicine but do not say what the truth-makers of their causal claims are.

BIBLIOGRAPHY

R. Amundson (2000) Against Normal Function, *Studies in History and Philosophy of Biological and Biomedical Sciences* 31, pp. 33–53.

R. Amundson, G. Lauder (1994) Function without Purpose: the Uses of Causal Role Function in Evolutionary Biology. *Biology and Philosophy* 9, pp. 443-469

E. de Angelis (2003) Causation in the Seventeenth Century, Final Causes, in *Dictionary of the History of Ideas*, (http://etext.virginia.edu/, accessed 09/02/2010)

A. Ariew (2002) Platonic and Aristotelian Roots of Teleological Arguments, in Ariew *et. al.* (eds.), Functions: New Reading in the Philosophy of Psychology and Biology, Oxford University Press (cited, online version, at http://web.missouri.edu/~ariewa/Teleology.pdf , last accessed 09/04/2010)

A. Ariew (2007) Teleology, in M. Ruse and D. Hull eds., *Cambridge Companion to the Philosophy of Biology*, Cambridge University Press, (cited, online version at, www.missouri.edu/~ariewa/teleologyII.pdf, last accessed 09/04/2010)

D. Armstrong (1985) *What is a law of nature?* Cambridge University Press

D. Armstrong (1997) *A world of states of affairs*, Cambridge University Press

D. Armstrong (1999) Comment on Ellis, in *Causation and Laws of Nature*, H. Sankey *ed.*, pp. 35-39, Springer

D. Armstrong, A. Heathcote (1991) Causes and Laws, *Nous*, 15(1) pp. 63-74

M. Arnold (2008) Heart Failure, in *The Merk Manual for Healthcare Professionals*, (http://www.merck.com/mmpe/sec07/ch074/ch074a.html, last accessed 02/05/2010)

J. Ayers (1981) Locke versus Aristotle on Natural Kinds, *Journal of Philosophy*, 78, pp. 247-272

M. Ayers (1991) *Locke* (vol. 1), Routledge

R.S. Bahn, C.M. Dutton, N. Natt, W. Joba, C. Spitzweg, A.E. Heufelder (1998) Thyrotropin receptor expression in Graves' orbital adipose/connective tissues: potential auto-antigen in Graves' ophthalmopathy, *Journal of Clinical Endocrinology and Metabolism*. 83, pp. 998-1002

H. Beebee (2001) Recent Work on Causation, *Philosophical Books* 42, pp. 33-45

J. Benjamins (2003) Review of Psychological Concepts and Biological Psychiatry: A Philosophical Analysis, *Psychological Medicine*, 33

J. Bigelow, R. Pargetter (1990) *Science and Necessity*, Cambridge University Press

J. Bigelow (1999) Scientific Ellisianism, in *Causation and Laws of Nature*, H. Sankey *ed.*, Springer

A. Bird (1998) *Philosophy of Science*, Routledge

A. Bird (1998) Dispositions and Antidotes, *Philosophical Quarterly* 48, pp. 227–34

A. Bird (2004) Antidotes All the Way Down?, *Theoria*, 19 (51), pp. 259-269

A. Bird (2005) Laws and Essences, *Ratio*, 18 (4), pp. 437–461

A. Bird (2007) *Nature's Metaphysics: Dispositions, Laws, and Properties*, Oxford University Press

A. Bird (2008) Causal Exclusion and Evolved Emergent Properties, in *Revitalizing Causality: Realism about causality in philosophy and social sciences*, R. Groff *ed.*, Routledge

A. Bird (2009) Kripke, in Twelve Modern Philosophers, C. Belshaw and G. Lemps (eds), Wiley-Blackwell

A. Bird, E. Tobin (2008) Natural Kinds, in *The Stanford Encyclopedia of Philosophy*, E. Zalta *(ed.)* (http://plato.stanford.edu/entries/natural-kinds/, last accessed 98/06/2010)

D. Bohm *et. al.* (1955) A Causal Interpretation of the Pauli Equation, *Il Nuovo Cimento*, 1 (1) pp. 48-66

C. Boorse (1977) Health as a Theoretical Concept, *Philosophy of Science*, 44, pp. 542–573.

G. Botterill (2005) Review of Ellis: Scientific Essentialism, *Philosophical Books*, Volume 46 Issue 2, pp. 118 – 122
R. Boyd (1991) Realism, Anti-Foundationalism, and the Enthusiasm for Natural Kinds, *Philosophical Studies* 61, pp. 127-148.
R. Boyd (1995) Observation Explanatory Power and Simplicity – Towards a non-Humean account in *The Philosophy of Science*, ed. Richard Boyd, pp. 349-376, MIT Press
R. Boyd (1999) Homeostasis, Species and Higher Taxa, in R. Wilson (ed.) *Species: new interdisciplinary essays*, pp. 141-185, MIT Press
G. Brent (2008) Graves' disease, *New England Journal of Medicine*, 358, pp. 2594-2605
J. Brigandt (2009) Natural Kinds in Evolution and Systematics: Metaphysical and Epistemological Considerations, *Acta Biotheoretica*, 57, pp. 77-93
B.A. Brody (1974) Why Settle for Anything Less Than Good Old-fashioned Aristotelian Essentialism? *Noûs* 7, pp. 351-365
A. Burgun *et al* (2005) Issues in the Classification of Disease Instances with Ontologies, *Stud Health Technol Inform*, 116, pp. 695–700
J. Canfield (1964) Teleological Explanation in Biology: *British Journal for the Philosophy of Science* 14, pp. 285-95
G. Canguilhem (1991) *The Normal and the Pathological*, trans. Carolyn R. Fawcett & Robert S. Cohen, New York: Zone Books, originally published as G. Canguilhem (1966) *Le normal et le pathologique, augmenté de Nouvelles réflexions concernant le normal et le pathologique*, PUF
R. Carnap (1995) *An Introduction to the philosophy of science*, Dover
N. Cartwright, J. Dupre (1988) Probability and Causality: Why Hume and Indeterminism Don't Mix, *Nous*, 22, pp. 521-36
N. Cartwright (1989) *Nature's Capacities and their Measurement*, Oxford University Press
N. Cartwright (1999) *The Dappled Word*, Cambridge University Press
N. Carwright (2001) Modularity : It Can-and Generally Does-Fail, in *Stochastic Causality*, Maria Carla Galavotti, Patrick Suppes and Domenico Costantini (eds.), CSLI Publications, , 65–84
N. Cartwright (2002a) From Causation to Explanation and Back, (http://personal.lse.ac.uk/cartwrig/Default.htm, last accessed 03/02/2010)
Cartwright (2002b) In Favour of Laws that are Not Ceteris Paribus After All, *Erkenntnis*, 57, pp. 425-439
N. Cartwright (2002c) 'Introduction' and 'Reply' in *Philosophical Books* (43), pp 241-244; 271-279
N. Cartwright (2005) 'How Can We Know What Made the Ratman Sick? Singular Causes and Population Probabilities', in *Philosophy of Religion, Physics, and Psychology: Essays in Honor of Adolf Grünbaum*, Aleksandar Jokić (ed.), Prometheus Books. (http://personal.lse.ac.uk/cartwrig/Default.htm, accessed 08/06/2010)
N. Cartwright (2007a) *Causal Powers*, Monograph LSE, (http://personal.lse.ac.uk/cartwrig/Default.htm, accessed 08/06/2010)
N. Cartwright (2007b) *Hunting Causes and Using Them*, Cambridge University Press
N. Cartwright, N. Alexandrova (2007) Laws in *The Oxford Handbook of Contemporary Philosophy, eds.* Frank Jackson, Michael Smith, Oxford University Press
B. Chabner, E. Chabner-Thompson (2008) Paraneoplastic Syndromes, in *The Merk Manual for Healthcare Professionals*, (http://www.merck.com/mmpe/sec11/ch147/ch147g.html, accessed 07/06/2010)
M. Chaplin (2008) *Water Structure and Science*, (http://www.lsbu.ac.uk/water/index2.html, last accessed 02/03/2010)
R. Chisholm (1967) Identity through Possible Worlds: Some Questions, *Nous*, 1 (1), pp. 1-8

L. Clapp (2002) Review of Brian Ellis's Scientific Essentialism, *The Philosophical Review* 111(4), 589—594
J. Collier (1996) On the Necessity of Natural Kinds, in Peter Riggs, ed. *Natural Kinds, Laws of Nature and Scientific Reasoning,* Kluwer
M. Colyvan (1998) Can the Eleatic Principle be Justified? *Canadian Journal of Philosophy* 28, pp. 313–336
R. Cooper (2002) Disease, *Studies in the History and Philosophy of Biology & the Biomedical Sciences*, 33: 263–282.
R. Cooper (2004) What's Wrong with DSM?, *History of Psychiatry*, 15(1): 005–025
R. Cooper (2005) *Classifying Madness*, Springer
R. Cooper (2008) Are There Natural Kinds in Psychology?, in *Revitalizing Causality: Realism about causality in philosophy and social sciences,* R. Groff *ed.*, Routledge
J. Crane (2003) Locke's Theory of Classification, *British Journal for the History of Philosophy*, 11(2), pp. 249-259
R. Cummins (1975) Functional Analysis, *The Journal of Philosophy* 72 (20) pp. 741-760.
W. Cushing (1993) Bohm's Theory: Common Sense Dismissed, *Studies In History and Philosophy of Science,* Part A 24 (5) pp. 815-842
C. Craver (2001) Role functions, mechanisms and hierarchy. *Philosophy of Science,* 68, pp. 31–55.
N. Dahl (1997) Two Kinds of Essence in Aristotle, *The Philosophical Review*, 106 (2), pp. 233-265
C. Daly (1996) Defending Promiscuous Realism, *Philosophical Quarterly*, 46(184), pp. 496-500
T. F. Davies, E. Roti, L. E. Braverman, L.I. DeGroot (1998) Thyroid controversy -- stimulating antibodies. *J Clin Endocrinol Metab,* 83, pp. 3777-3785.
A. Denkel (1996) *Object and Property*, Cambridge University Press
K. Donnellan (1973) Substances as Individuals, *The Journal of Philosophy*, 70 (19) pp. 711-712
A. Drewery (2001) Dispositions and Ceteris Paribus Laws, *British Journal for the Philosophy of Science*, 52 (4) pp. 723-733
J. Dupré (1981) Natural Kinds and Biological Taxa, *Philosophical Review,* 90, pp. 66-91
J. Dupré (1993) *The Disorder of Things,* Harvard University Press
J. Dupré, (2001) In Defence of Classification, *Studies in the History and Philosophy of the Biological and Biomedical Sciences* 32: 203–219; reprinted in *Humans and Other Animals* (2002), pp. 81–99, Clarendon
J. Dupré, (2002) *Humans and other Animals*, Oxford University Press.
J. Earman, J. Roberts, and S. Smith (2002) '*Ceteris Paribus*' Lost", *Erkenntnis,* 57 (3), pp. 281–301
C. L. Elder (2005) Undercutting the Idea of Carving Reality, *Southern Journal of Philosophy,* 43, pp. 41-59.
C. L. Elder (2007) Realism and the Problem of Infimae Species, *American Philosophical Quarterly*, 44: 111-27
B. Ellis *et. al.* (1992) The World as One of a Kind: Natural Necessity and Laws of Nature, *The British Journal for the Philosophy of Science*, 43 (3) pp. 371-88
B. Ellis (1999) Bigelow's Worries about Scientific Essentialism, in *Causation and Laws of Nature*, ed. Howard Sankey, Springer
B. Ellis (2000) Causal Laws and Singular Causation, *Philosophy and Phenomenological Research*, 61 (2), pp. 329-351
B. Ellis (2001) *Scientific Essentialism*, Cambridge University Press
B. Ellis (2002) *The Philosophy of Nature*, McGill-Queen's University Press
B. Ellis (2005a) Universals, the Essential Problem and Categorical Properties, *Ratio*, 18 (4), pp. 462-472
B. Ellis (2005b) Physical Realism, *Ratio*, 18 (4) pp. 371-384
B. Ellis (2008) Essentialism and Natural Kinds, in S. Psillos and M. Curd (eds.), *Routledge Companion to Philosophy of Science,* pp 139-148, Routledge

M. Ereshefsky (1998) Species Pluralism and Anti-Realism, *Philosophy of Science*, 65:103-120
M. Ereshefsky (2002) Linnaean Hierarchy: Vestiges of a Bygone Era, *Philosophy of Science* 69, pp. S305-S315
E. Fales (1990) *Causation and Universals,* Routledge
R. Fenyman *et. al* (1966) *The Feynman Lectures on Physics* (vol. 3 : quantum mechanics), Addison-Wesley
M. Fernet *et. al.* (1987) A case of ectopic right renal artery: a radiologic-anatomic variant, *Surgical and Radiologic Anatomy,* 9(4), pp. 319-20.
J. A. Fodor (1974) Special sciences (or: The disunity of science as a working hypothesis) *Synthese*, 28 (2) pp. 97-115
J. Foster (2004) *The Divine Lawmaker,* Oxford University Press
K W M Fulford (2001) What is (mental) disease?': an open letter to Christopher Boorse, *J Med Ethics 27,* pp. 80-85
A. Funkenstein (1989) *Theology and Scientific Imagination,* Princeton University Press
P. Geach (1962) *Reference and Generality,* Ithaca
R. Giere (2000) Review of 'Dappled World' in *Philosophy of Science,* 67 (3) pp. 527-530
J.H. Gillard *et. al* (1998) Riedel's Lobe of the Liver: Fact or Fiction? *Clinical Anatomy* 11:47–49
J. Ginsberg (2003) Diagnosis and management of Graves' disease, *Canadian Medical Association Journal,* 168 (5) pp. 575-586
P. Godfrey-Smith (1993) Functions: Consensus Without Unity (cited, online version, at http://www.people.fas.harvard.edu/~pgs/OnlinePapers/PGSFunctions1993PPQ.pdf, last accessed 04/03/2010)
P. Godfrey-Smith (1994) A Modern History Theory of Functions, *Noûs* 28 (3), pp. 344-362
N. Goodman (1946) A Query on Confirmation, *The Journal of Philosophy,* 43 (14), pp. 383-385
N. Goodman (1965) *Fact, Fiction, and Forecast,* Bobbs-Merrill
R. Grady (2007) Sortals, in *The Stanford Encyclopaedia of Philosophy,* E. Zalta *(ed.),* (http://plato.stanford.edu/entries/sortals/, last accessed 05/02/2010)
P.E. Griffiths (1997) *What emotions really are: The problem of psychological categories,* University of Chicago Press
I. Hacking (1984) *The Emergence of Probability,* Cambridge University Press
I. Hacking (1993) Working in a New World: The Taxonomic Solution, in Horwich *ed., World Changes,* MIT Press
I. Hacking (2006) A Tradition of Natural Kinds, pp. 370-387, in *Philosophy of Science: An Anthology,* M. Lange *ed.,* Blackwell Publishing
I. Hacking (2007) Natural Kinds: Rosy Dawn, Scholastic Twilight, *Royal Institute of Philosophy Supplement,* 82, pp. 203-239
S. Hampshire (2005) *Spinoza and Spinozism,* Oxford University Press
D. M. Hausman, J. Woodward (2004) Modularity and the Causal Markov Condition, *British Journal for the Philosophy of Science,* 55 (1) pp. 147-161
T. Handfield (2009) *Dispositions and Causes,* Oxford University Press
R. Harré, *R.* (2005) Chemical *Kinds* and Essences Revisited, *Foundations of Chemistry.* 7(1) pp. 7--30
N. Haslam (2002) Natural kinds, Practical Kinds and Psychiatric Categories, *Psycoloquy:* 13(001)
N. Haslam (2003) Essentialist Beliefs about Mental Disorders, *Journal of Social and Clinical Psychology,* 21 (6) pp. 628-644
J. Heil (2006) *Universals,* paper delivered at the Lowe conference, Buffalo University, (http://ontology.buffalo.edu/06/Lowe/Heil.pdf, last accessed 01/02/2010)

L. Hemming (2000) After Heidegger – Transubstantiation, *Heythrop Journal*, 42(2), pp. 170–186
C. Hempel (1966) *Philosophy of Natural Science*, Prentice-Hall
E. Hirsch (1982) *The Concept of Identity*, Oxford University Press
E. Hirsch (1997) *Dividing Reality*, Oxford University Press
L. E. Holm (2005) Smoking and other lifestyle factors and the risk of Graves' hyperthyroidis, *Archives of Internal Medicine*, 165 (14) pp. 1606-1611
P. Hylton (2007) *Quine*, Routledge
A. Koyre (1957) *From the Closed World to the Infinite Universe*, John Hopkins University Press
G. Irzik (1992) Cartwright, Capacities and Probabilities, *PSA*, 1, pp. 239-250
D. L. Jacobson, S. J. Gange, N. R. Rose, N. M. Graham (1997) Epidemiology and estimated population burden of selected autoimmune diseases in the United States, *Clin Immunol Immunopathol*, 84, pp. 223-243
E. M. Jacobson, Y. Tomer (2007) The Genetic Basis of Thyroid Autoimmunity, *Thyroid*, 17, pp. 949-961
G. Janelle (2003) Valvular pathology diagnosed with transesophageal echocardiography during aortic root replacement, *Journal of Cardiothoracic and Vascular Anesthesia*, 17 (2), pp. 271-272
E. Juengst (1983) Causation and the Conceptual Scheme of Medical Knowledge, in *Science, technology, and the art of medicine: European-American dialogue*, Corinna Delkeskamp-Hayes, Mary Ann Gardell Cutter *(eds.)*, pp. 127-153, Springer
Kasper DL, Braunwald E, Fauci AS, Hauser SL, Longo DL, Jameson JL, Loscalzo J. (2008) *Harrison's principles of internal medicine* (17th ed.), McGraw-Hill
M. Khalidi (1998) Natural Kinds and Crosscutting Categories, *The Journal of Philosophy*, 95 (1), pp. 33-50
M. Khalidi (1993) Carving Nature at the Joints, *Philosophy of Science* 60, pp. 100-113
P. Kitcher (1993) Function and Design, *Midwest Studies in Philosophy* 28, pp. 379–397
G. Klima (2008) The Medieval Problem of Universals, Stanford Encyclopedia of Philosophy. E. Zalta ed. (http://plato.stanford.edu/entries/universals-medieval/ accessed 03/05/2010)
H. Kornblith (1995) *Inductive Inference and Its Natural Ground: An Essay in Naturalistic Epistemology*, MIT Press
S. Kripke (1980) *Naming and Necessity*, Wiley-Blackwell
H. Lagerlund (2009) John Buridan's Empiricism and the Knowledge of Substances (http://faculty.fordham.edu/klima/FILES/Talk%20Western%20Ontario%20Oct%20-09-revised.pdf, accessed 3/05/2010)
M. Lange (2000) *Natural Laws in Scientific Practice*, Oxford University Press
M. Lange (2002) Who's afraid of ceteris-paribus laws? Or: how I learned to stop worrying and love them" *Erkenntnis*, 57, pp. 407-423
M. Lange (2004) A Note on Scientific Essentialism, Laws of Nature, and Counterfactual Conditionals, *Australasian Journal of Philosophy* 82 (2), pp. 227 – 241
J. Laporte (2004) *Natural Kinds and Conceptual Change*, Cambridge University Press
D. Lewis (1986) Causal Explanation, in *Philosophical Papers, Volume II*, pp. 214–240, Oxford University Press
R. Lipton, T. Ødegaard (2005) Causal thinking and causal language in epidemiology: it's in the details *Epidemiologic Perspectives & Innovations*, 2, p. 8
C. J. Lote (2000) *Principles of Renal Physiology*, Kluwer
J. Lowe (1989) *Kinds of Being*, Blackwell
J. Lowe (1995) Coinciding Objects: in Defence of the Standard Account, *Analysis*, 55: 171–78
J. Lowe (1998) Entity, Identity and Unity, *Erkenntnis* 48: 191–208
J. Lowe (2000) *An Introduction to the Philosophy of Mind*, Cambridge University Press

J. Lowe (2001) *The Possibility of Metaphysics*, Oxford University Press
J. Lowe (2003) Review: Sameness and Substance Renewed, *Mind*, 112, pp. 816-820
J. Lowe (2006) *The Four Category Ontology*, Oxford University Press
J. Lowe (2007a) A Problem for A Posteriori Essentialism Concerning Natural Kinds. *Analysis* 67(296), pp. 286-292
J. Lowe (2007b) Essentialism and Chemical Kinds, *work in progress paper*, (http://www.bris.ac.uk/metaphysicsofscience/workinprogress/loweesschem.pdf, last accessed 03/02/2010)
J. Lowe (2008). Two Notions of Being: Entity and Essence. *Royal Institute of Philosophy Supplements* 83 (62), pp. 23-48
J. Lowe (2009) Individuation, in *A Companion to Metaphysics*, 2nd edition, J. Kim, E. Sosa and G. Rosenkrantz *eds.*, pp. 28-36, Blackwell
E. Mackery (2005) Concepts Are Not a Natural Kind, *Philosophy of Science*, 72, pp. 444–467
P. Mackie (1994) Sortal Concepts and Essential Properties, *The Philosophical Quarterly*, 44 (176) pp. 311-333
P. Mackie (2006) Trans-world Identity, in *The Stanford Encyclopedia of Philosophy*, E. Zalta *(ed.)*, (http://plato.stanford.edu/entries/identity-transworld/, last accessed 03/02/2010)
N. Manji, J. D. Carr-Smith, K. Boelaert K, *et al* (2006) Influences of age, gender, smoking, and family history on autoimmune thyroid disease phenotype, *J Clin Endocrinol Metab*, 91, pp. 4873-4880
M. C. Many, S. Costagliola, M. Detrait, J-F. Denef, G. Vassart, and M. Ludgate. (1999) Development of an animal model of autoimmune thyroid eye disease. *Journal of Immunology* 162, pp. 4966-4974
R. B. Marcus (1993) *Modalities: Philosophical Essays*, Oxford University Press
J. Marcum (2008) *Humanizing Modern Medicine: An Introductory Philosophy of Medicine*, Springer
J. Margolis (1976) The concept of disease, *The Journal of Medicine and Philosophy* 1(3), pp. 238-255
R. J. McNally (2000) On *Wakefield's* harmful dysfunction analysis of mental disorder, *Behaviour Research and Therapy*, pp. 309–314.
C. Megone (1998) Aristotle's Function Argument and the Concept of Mental Illness *Philosophy, Psychiatry, & Psychology*, 5 (3) pp. 187-201
H. Mellor (1977) Natural Kinds, *The British Journal for the Philosophy of Science*, 28(4), pp. 299-312
P. Menzies (2002) Capacities, Natures and Pluralism: A New Metaphysics for Science, *Philosophical Books*, (43) pp.261-270 (http://www.phil.mq.edu.au/staff/pmenzies/cartwright.PDF , last accessed 03/02/2010)
R. G. Millikan, (1999) Historical Kinds and the "Special Sciences", *Philosophical Studies*, 95 (1/2) pp. 45-65
S. Mitchell (2000) Dimensions of Scientific Law, *Philosophy of Science*, 67 (2) pp. 242-265
W. E. Morris (2007) David Hume, in *The Stanford Encyclopedia of Philosophy*, E. Zalta *(ed.)*, http://plato.stanford.edu/entries/hume/, last accessed 03/02/2010)
M. Morrison (1995), Capacities, Tendencies and the Problem of Singular Causes, *Philosophy and Phenomenological Research* 55(1) pp. 163-168
M. P. Mourits, M. F. Prummel, W. M. Wiersinga, L. Koornneef (1997) Clinical activity score as a guide in the management of patients with Graves' ophthalmopathy, *Clin Endocrinol*, 47, pp. 9-14
S. Mumford (1998) *Dispositions*, Oxford University Press
S. Mumford (2000) Normative and Natural Laws, *Philosophy*, 75, pp. 265-282
S. Mumford (2004) *Laws in Nature*, Routledge
S. Mumford (2005) Kinds, Essences, Powers, *Ratio* 18, pp. 420-436
D. Murphy (2006) *Psychiatry in the Scientific Image*, MIT Press

E. Nagel (1961), *The Structure of Science: Problems in the Logic of Scientific Explanation*, Brace & World
S. Nandler (1998) Doctrines of explanation on late scholasticism and in the mechanical philosophy, in *The Cambridge History of the Seventeenth Century Philosophy*, eds. D Garber and M. Ayers, pp. 513-532, Cambridge University Press
K. Neander (1991) The Teleological Notion of Function, *Australasian Journal of Philosophy*, 69 (4), pp. 54-68
A. Nelson (1990) Are Economic Kinds Natural? in *Scientific Theories*, ed. C. Wade Savage, pp. 102-134, University of Minnesota Press
R. M. Nesse, G. C. Williams (1999) Research designs that address evolutionary questions about medical disorders. In S. Stearns (Ed.), *Evolution in Health and Disease* pp. 16-26, Oxford University Press
B. Nickel (2008) Natural Kinds and Induction in the Special Sciences, *work in progress paper*, (www.bernhardnickel.net/papers/Spec-Sci-Induction.pdf , last accessed 03/02/2010)
H. Noonan (1981) Reviewed work(s): Sameness and Substance. by David Wiggins, *The Philosophical Quarterly*, 31, (124) pp. 260-268
H. Noonan (2006) Identity, in *The Stanford Encyclopaedia of Philosophy*, E. Zalta *(ed.)*, (http://plato.stanford.edu/entries/identity/, last accessed 03/02/2010)
D.S. Oderberg (1996) Coincidence under a Sortal, *The Philosophical Review*, 105 (2) pp. 145-171.
D.S. Oderberg (2005) Hylermorphic dualism, in *Personal identity*, eds. Ellen Frankel Paul, Fred Dycus Miller, Jeffrey Paul, Cambridge University Press
D.S. Oderberg (2007) *Real Essentialism*, Routledge
D. Papineau (1991) Review: Correlations and Causes,*The British Journal for the Philosophy of Science*, 42 (3) pp. 397-412
L.A. Paul (2002) Limited Realism: Cartwright on Natures and Laws, *Philosophical Books*,(cited, online version at http://www.u.arizona.edu/~lapaul/papers/NC-review.pdf , last accessed 23/01/2010)
W. Pauli (1945) *Exclusion Principle and Quantum Mechanics*, Nobel Lecture, (http://nobelprize.org/nobel_prizes/physics/laureates/1945/pauli-lecture.pdf, last accessed 03/04/2010)
S. Psillos (2002) *Causation and Explanation*, Acumen
S. Psillos (2007) *Philosophy of Science A-Z*, Edinburgh University Press Ltd
S. Psillos (2008) Cartwright's Realist Toil, pp.167-194 in *Cartwright's Philosophy of Science*, (eds.) Stephan Hartman, Carl Hoefer & Luc Bovens, Routledge (http://www.phs.uoa.gr/~psillos/Publications_files/Cartwright-Psillos.pdf, last accessed 15/03/2010)
W. Quine (1968) Natural Kinds in *Ontological Relativity and Other Essays*, Columbia University Press
L. Reznek (1987) *The Nature of Disease*, Routledge & Kegan Paul
F. Russo and J. Williamson (2007) Interpreting Causality in the Health Sciences, *International Studies in the Philosophy of Science* 21(2), pp. 157-170. (online, cited version at http://www.kent.ac.uk/secl/philosophy/jw/2006/interpreting_causality.pdf, last accessed 22/02/2010)
B. Rapoport, G. D. Chazenbalk, J. C. Jaume, S. M. McLachlan (1998) The thyrotropin (TSH) receptor: interaction with TSH and autoantibodies, *Endocr Rev*, 19, pp. 673-716.
H. Robinson (2004), Substance, in *The Stanford Encyclopaedia of Philosophy*, E. Zalta *(ed.)*, (http://plato.stanford.edu/entries/substance/, last accessed 22/02/2010)
J. Sadler, G. Agich (1995) Diseases, Functions, Values, and Psychiatric Classification, in *Philosophy, Psychiatry, & Psychology*, 2 (3) pp. 219-231
N. Salmon (1979) How Not to Derive Essentialism From the Theory of Reference, *The Journal of Philosophy*, 703—725
N. Salmon (1982) *Reference and Essence*, Blackwell
W. Salmon (1998) *Causality and Explanation*, Oxford University Press

K. Schaffner (1993) *Discovery and Explanation in Biology and Medicine*, Chicago University Press

J. Schaffer (2007) The Metaphysics of Causation, in *The Stanford Encyclopaedia of Philosophy*, E. Zalta *(ed.)*, (http://plato.stanford.edu/entries/causation-metaphysics/, last accessed 03/02/2010)

J. Schaffer (2008) Causation and Laws of Nature: Reductionism, in *Contemporary Debates in Metaphysics*, T. Seider, J. Hawthrone and D. Zimmerman *(eds.)*, Blackwell

M. Schrenk (2007) Can Capacities Rescue Us From *Ceteris Paribus* Laws? In B. Gnassounou & M. Kistler (eds.), *Dispositions in Philosophy and Science*, Ashgate.

M. Schrenk (2008) *A Theory for Special Science Laws* in Walter, S., et al. (eds.) Selected Papers Contributed to the Sections of GAP.6, 6th International Congress of the Society for Analytical Philosophy, Berlin

R. Shain (1993) Mill, Quine, and Natural Kinds, *Metaphilosophy*, 24, pp. 275-292

N. Sharpe, K. Swedgerg (1999) Angiotensin Converting Inhibitors, in *Heart Failure Management*, N. Shape (ed.) Taylor & Francis

S. Shoemaker (1970) Wiggins on Identity, *The Philosophical Review*, 79 (4) pp. 529-544

S. Shoemaker (1998) Causal and Metaphysical Necessity, *Pacific Philosophical Quarterly*, 79 (1), pp. 59-77

A. Sidelle (1992) Identity and the Identity-like, *Philosophical Topics* 20, pp. 269-292

T. Sider (2006) Bare Particulars, *Philosophical Perspectives* 20. Pp/ 387-97

L.J. Snyder (2006) Whewell, in *The Stanford Encyclopaedia of Philosophy, E. Zalta (ed.)*, http://plato.stanford.edu/entries/whewell/#SciInd

S. Soames (2006) The Philosophical Significance of The Kripkean Necessary Aposteriory, in *Philosophical Issues*, Vol. 16, E Sosa and E. Vilaneuva *(eds.)*, Blackwell Publishing http://www-rcf.usc.edu/~soames/sel_pub/Kripkean_Necessary_Aposteriori.pdf

E. Sober (1999) The Multiple Realizability Argument against Reductionism, *Philosophy of Science* 66, pp. 542–564

H. Stewart (1990) Identity Statements and the Necessary A Posteriori, *The Journal of Philosophy*, 87 (8) pp. 385-398

D. Sweet (2007) *Heraclitus: Translation and Analysis*, Rowman & Littlefield

D. Sulmasy (2005) Diseases and Natural Kinds, *Theoretical Medicine and Bioethics* 26, pp. 487–513

P. Suppes (1970) *A Probabilistic Theory of Causality*, North Holland

T. Szasz (1987) *Insanity*, Wiley

P. Tanser (2007) Aortic Regurgitation, in *The Merk Manual for Healthcare Professionals*, (http://www.merck.com/mmpe/sec07/ch076/ch076b.html, accessed 04/04/2010)

P. Teller (1975) Essential Properties: Some Problems and Conjectures, *The Journal of Philosophy*, 72(9), pp. 233-248

M. Thalos (1995) Review of 'The Disorder of Things', *Philosophy of Science*, 62 (2), pp. 351-355

H. Thomson (1969) Species, Determinates and Natural Kinds, *Nous* 3, pp. 95-101

E. Tobin (2008) Natural Kinds and Symbiosis, *work in progress paper*, (http://www.bris.ac.uk/metaphysicsofscience/workinprogress/nksymbiosis.pdf last accessed 03/02/2010)

J. Tsou (2008) *The Reality and Classification of Mental Disorders*, PhD thesis, University of Chicago http://home.uchicago.edu/~jtsou/Tsou_Dissertation_2008.pdf (last accessed 03/02/2010)

W. Uzgalis (1988) The Anti-Essentialist Locke and Natural Kinds, *The Philosophical Quarterly* 38 (152), pp. 330-339

B. Van Fraassen (1980) *The Scientific Image*, Oxford University Press

B. Van Fraassen (1989) *Laws and Symmetry*, Oxford University Press

J. Wakefield (1992) The Concept of Mental Disorder, *American Psychologist*, 47, pp. 373–388

J. Wakefield (2001) Evolutionary history versus current causal role in the definition of disorder: reply to McNally, *Behaviour Research and Therapy*, 39 (3) pp. 347-366

G. Wallerstein, I. Iben, P. Parker, *et. al.* (1997) Synthesis of the elements in stars: Forty years of progress. *Reviews of Modern Physics* 69(4) (http://cococubed.asu.edu/papers/wallersteing7.pdf, last accessed 03/02/2010)

D.M. Walsh (2008) Functions, in *Routledge Companion to Philosophy of Science*, S. Psillos and M. Curd (*eds*), Routledge

B. Weatherson (2002) Review of David Wiggins, 'Sameness and Substance Renewed', *Notre Dame Philosophical Reviews*

A. P. Wheetman (2000) Graves' disease. *New England Journal of Medicine* 343(17), pp. 1236-1248

C. Whitbeck (1977) *Causation in Medicine: The Disease Entity Model, Philosophy of Science* 44 (4), pp. 619-637

D. Wiggins (1967) *Identity and Spatio-Temporal Continuity*, Blackwell

D. Wiggins (1980) *Sameness and Substance*, Cambridge University Press

D. Wiggins (2001) *Sameness and Substance Renewed*, Cambridge University Press

T. Wilkerson (1995) *Natural Kinds*, Ashgate Publishing Ltd

R. Willenheimer, B. Dahlöf, E. Rydberg, L. Erhardt, (1999) AT_1-receptor blockers in hypertension and heart failure: clinical experience and future directions, *Eur Heart J*, 20, pp. 997-1008

J. Wilson (1999) *Biological Individuality: The Identity and Persistence of Living Entities*, Cambridge University Press

J. Wilson (2005), *Causality*, (cited, online version at http://individual.utoronto.ca/jnwilson/E_C.pdf, last accessed 03/02/2010)

R.A. Wilson (1996) Promiscuous Realism, *British Journal for the Philosophy of Science*, 47 (2), pp. 303-316

A. Wouters (1999) *Explanation Without A Cause*. Ph.D. thesis, Utrecht University (http://www.morepork.demon.nl/diss/index.html, last accessed 03/02/2010)

A. Wouters (2005) The Function Debate in Philosophy, *Acta Biotheoretica* 53(2), pp. 123-15

J. Woodward (2002) There is no such Thing as a ceteris paribus Law, *Erkenntnis*, 57(3) pp. 303–328

J. Woodword (2008) Causation and Manipulability, in *The Stanford Encyclopaedia of Philosophy*, E. Zalta *(ed.)*, http://plato.stanford.edu/entries/causation-mani/, last accessed 21/04/2010)

L. Wright (1973) Functions, *Philosophical Review*, 82(2), pp. 139-168

F. Wuketits (1984) *Concepts and Approaches in Evolutionary Epistemology*, Springer

S. Yablo (2003) "Tables Shmables," review of D. Wiggins, *Sameness and Substance Renewed*, in *Times Literary Supplement*, (www.mit.edu/~yablo/wiggins.pdf, last accessed 03/06/2010)

P. Zachar (2000) Psychiatric Disorders Are Not Natural Kinds, *Philosophy, Psychiatry, & Psychology*, 7 (3), pp. 167-182