CHAPTER ONE: INTRODUCTION

In the most general terms, causal realism is the philosophical doctrine according to which causality is a real, objective feature of the world. It is contrasted to causal anti-realism, which holds that causality is an aspect of our apprehension of the world, not of the world itself. This thesis is a defense of the plausibility and coherence of causal realism.

Contemporary versions of causal anti-realism have in common the thesis that causality refers to the regularity with which phenomena are observed to follow certain others. These approaches bear the obvious influence of David Hume, whose theory of causality was developed in *A Treatise of Human Nature* (1739) and *An Enquiry Concerning Human Understanding* (1748). The historical origins of the philosophical study of causality predate Hume's writings by some two millennia, however.

1.1 The Classical Roots of Causal Realism

The argument for causal realism is, as I frame it, informed at several points by assumptions deriving from important parts of that history, *i.e.*, of what might broadly be characterized as the "Aristotelian" tradition. First, we need to examine the historical context in which Aristotle's metaphysical views were formulated, to understand the terms in which the concept of causality was originally understood, and to understand Aristotle's contribution to the debates of classical antiquity. An appreciation of Aristotle will allow us to note points of similarity with the doctrines of John Locke, the most Aristotelian of the pre-Humean moderns. Insofar as Hume was reacting to the implicit Aristotelianism in Locke's thought, an understanding of key parts of Aristotle's and Locke's metaphysics will help sharpen the contrast between Hume's ontologically conservative brand of empiricism and the more liberal brand of his Aristotelian predecessors.

Various aspects of causal realism have been defended this century in a more or less piecemeal fashion by H.W.B. Joseph, C. J. Ducasse, Sterling Lamprecht, E. H. Madden, Rom Harré, John Searle, Wesley Salmon, Nancy Cartwright, and James Woodward, among others. Some, such as Cartwright and Madden, are quite open about the Aristotelian and Lockean influences on their conception of causality. Given that they both hold that part of the case for causal realism is based on an assessment of the presuppositions of scientific reasoning, it might seem surprising that this is the case. Indeed it might seem anachronistic, and even bizarrely so. It is perhaps hard to see how an Aristotelian vision of science could have any contemporary philosophical relevance, given the advances in scientific method made during the Renaissance, the Enlightenment and the Victorian era.

One can easily grant that our understanding of the nature of science has improved tenfold or more over the last two millenia. The question is, how have those advances been achieved?—by *a priori* deduction from some set of self-evident truths or another? I submit that the methodological advances that have been made come from deep *a posteriori* reflection upon the nature of novel, progress-yielding scientific practices. By that, I mean that the novel practices in question improve (a) the accuracy and/or precision of our measurements, (b) the scope of their mathematical representation, (c) the range of manipulation and control that we have over nature, or (d) the reproducibility of experiments. There is no doubt that the Aristotelian vision of science in general has been eclipsed, but what of the metaphysics that underlies this vision? Suppose that the sort of "methodological *a posteriorism*" I endorse is valid. One could then more easily concede that the modifications of Aristotelian metaphysical doctrine required to maintain philosophical plausibility are responses to growth in *scientific* knowledge rather than significant progress in the philosophical

arena.

Locke's own work in metaphysics can be viewed, in part, as an attempt to integrate Aristotelian metaphysical notions with both Renaissance empiricism and the corpuscularian mechanism popular among 17th century British scientists. Similarly, the present account can be viewed as my attempt to integrate the tenable (in my estimation) parts of Aristotelian and Lockean metaphysics with a realist interpretation of the methodological presumptions of contemporary empirical science.

Aristotle's own metaphysics in general, and his theories of causality in particular, were formulated in part as a response to the debates among his contemporaries concerning the explanation of diversity, motion, and change. It was within the context of these early metaphysical controversies out of which the basic theoretical motifs and concepts relating to causal phenomena emerged, and in the wake of which these concepts acquired their first systematic treatments. We know that the concept of cause, or αἰτία, was probably used first in legal contexts, where it meant that to which blame or responsibility is to be attributed. But since my principal interest is in analyses addressing the role of causality in nature, I will focus my attention on the emergence of causal concepts in ancient Greek theories of nature.

§1. Persistence and Change

In classical antiquity, the concept of "Being" referred to the permanent or unchanging aspects of the world, thus Being was contrasted with change, or "Becoming." (In order to avoid ambiguity, I use the term "persistence" to refer to qualitative self-identity over time, or the stability-in-being of things.) In the pre-Socratic era, it was widely believed that there was a contradiction involved in claiming the real existence of both persistence and change. The appearance of contradiction between these two basic principles was made vivid by the contrast between two ancient metaphysical systems: that of Heracleitus on one hand, and that of the Eleatics (*esp*. Parmenides and Zeno) on the other. Ensuing philosophical thought was preoccupied for some time with the possibility of reconciliation. The elimination of the dichotomy between persistence and change remained an elusive goal until Aristotle elaborated a system of metaphysics sufficiently robust to do the job.

From a modern Humean point of view, it may be difficult to see of what relevance this quest for reconciliation is to the subject of causality. Clarification may be forthcoming if a different perspective is adopted. Consider the realist thesis that causality refers to the phenomenon of *particulars acting in accordance with their natures*. This implies three facts about particulars, two of which are obvious, one perhaps not so obvious. First, it implies that there are particulars—that numerical diversity is real. Second, it implies that these particulars act—that they can move or otherwise undergo changes. Third, it implies that particulars have a nature—that there is some qualitatively invariant aspect of any acting entity in virtue of which it makes sense to say that *it* undergoes the motions or changes ascribed to it. That is, if *some* thing is a cause, then some *thing* is a cause. If there is no thing which we can identify as the subject of predication in any statement of causality, it follows that the notion of causality itself is invalid. Since causal realism entails the reality of diversity, motion and change, the realist must be in a position to refute any competing theories which deny reality to any of them.

In order to see how the appearance of contradiction might arise, consider what it means to exist—to be. To be some *thing* (as opposed to being "nothing"), is to be *some* thing (as opposed to another thing). The concept of existence analytically entails εἶναι τι—being something or other, being something definite. The notion of existence-as-such, or existence for subjects of an indeterminate nature, is incoherent. Now, to *change* is to become something that is not altogether what

previously was. If to exist is to be something determinate, that means being determinate "all the way down," *i.e.,* in all respects. Therefore, for something to change would mean that that something became other than what it is, which is a contradiction. If the nature of change is incompatible with the nature of being, then either being is an illusion, and all is change, or change is an illusion, and all is simple, unchanging and unitary. The latter horn of the dilemma was seized by the Eleatics, while the Heracleiteans impaled themselves on the former.

Heracleitanism is the doctrine according to which the cosmos consists of the constant becoming and passing away of all things in an immutable, ceaseless march of the transitory. Motion and change are the manifest forms of this eternal process, whose basic principle or ἀρχή, is fire.[1] Fire is the pure form of motion or change and while motion is always present, it comes in various pedigrees. The other elements—earth, air and water—are, like fire, *modes of transition* and not any sort of self-identical substantial "stuff." Any apparent stability is transitory and fleeting; there is no persistence, no eternal unchanging "being." In the Heracleitan universe, everything is unstable. There are, however, two general tendencies inherent in the world, which Heracleitus expressed in the doctrine of the two "Ways."[2] The first Way is the change away from fire, to more dense forms—the densest being earth. The second Way involves the change into more rarified forms, the rarest being fire. Temporary dynamic equilibria are possible when these two tendencies are balanced, resulting in regions of pseudo-stability, creating the illusion of persisting things and qualities.

Eleatic metaphysics, in all its basic theses, is the opposite of Heracleitus'. Whereas Heracleitus denied the reality of persistence, maintaining that all is change, the Eleatics held that all is *being.* For Parmenides, Existence exists (ἔστιν εἶναι) and only *existence* exists. Since non-existence cannot be a possible content of thought, it is only existing beings which *are.* This was taken to imply that Existence only *exists*—that there is nothing else ultimately to be said about it save for the way in which it exists—as an eternal, unchangeable, undifferentiated unity. Furthermore, if only existence exists, it is incoherent to assert the existence of regions of nothingness in between existing things—the idea of empty space is unintelligible. Assuming that all change involves motion, it follows from Parmenidean assumptions that there can be no motion, and hence no change.

The difficulties facing these two theories is evident even in such brief sketches. The Eleatic theory is forced to reject all of the perceptual evidence for the reality of diversity, locomotion, decay, *etc.,* and designate all such evidence as illusory. The only aspects of reality of which we could have veridical perceptual experience are its unchanging, eternal aspects. The Heracleitan theory is incompatible with the idea of inherently stable substances having identities that persist over time. It implies that we can separate the notion of action from actor, and suppose that actions need not have subjects which perform them; the actions themselves are ontologically basic. The challenge of overcoming the being/change dichotomy is to formulate an ontology which accepts the reality of both and identifies what it is that persists, and what it is that is capable of change.

Parmenides had argued that the concepts of exnihilation and annihilation are incoherent. Non-being cannot spawn being, nor can being negate its own essence. But even those who accepted Parmenides' doctrine of "existential inertia" recognized that diversity, motion and change needed to be explained, not explained *away*. Empedocles, for example, believed that creation and destruction were explained as acts of combination and separation of simpler "elements," which were

---

[1] Wilhelm Windelband, *History of Ancient Philosophy*, trans. Herbert E. Cushman (New York: Dover, 1956), 52.

[2] Ibid., 55.

themselves unoriginated, imperishable and immutable substances. Empedocles basically took over Heracleitus' modes of transition and "materialized" them, giving them a characterization detailed enough (albeit fanciful and arbitrary) to entrench their usage. Now, since the elements were immutable and unchanging, they are entirely passive and incapable of originating their own action. There must then be some independently existing principle(s) of force or power in the world to explain the motion of the elements.

While Empedocles' immaterial forces were represented metaphorically, Anaxagoras attributed the cause of motion to a single kind of material: a rarified substance which has an inherent and irreducible capacity for self-movement. As the lightest of all elements (the total number of which Anaxagoras saw no good reason to limit to four) it is the only one which has dynamic potential. It is a vitalizing ether that animates the other passive elements, and material composites thereof. Its cosmological role is essentially the same as that of Heracleitus' fire. According to Anaxagoras, this ethereal vital matter (or νοῦς) causes it to move. But the mode of causality involved is more like teleological causality than like efficient causality. Anaxagoras apparently thought that this living matter had the capacity to organize the passive matter into which it became infused, and impel it in various ways, after which mechanical interactions could sustain the motion of passive bodies. The νοῦς therefore explains both the existence of motion in general and the organization of matter into purposeful complexes whose configuration could not be explained on the basis of mechanical processes alone.

The atomism of Leucippus (like Democritus after him) represents a return to a more Parmenidean ontology. The atomists retained the conception of that which *is* as fundamentally homogeneous, indestructible and indivisible. The essential difference is that the Parmenidean plenum is broken up into particulate fragments (ἄτομοι), separated by empty space. The human experience of the cosmos is a product of the motion, collision, and aggregation of these atoms, which differ from one another only with respect to the properties which exhaustively characterize each of them—determinate measures of shape, size, and weight. Qualitative diversity, motion, change, as well as apparent creation and destruction, are all explained in terms of differences in quantitative arrangements of atoms, while the properties of the atoms themselves are absolutely stable. Empedocles and Anaxagoras believed that the passivity of the elements was such that an additional external factor was needed in order to explain the motion of things. The atomists, on the other hand, ascribed a random motion to atoms as part of their essential nature; they did not see any reason why motion in general could not be treated as an "immanent quality."[3] The posit of some ethereal vitalizing force is a theoretical redundancy in the eyes of the atomist. All diversity, motion and change can be accounted for in quantitative terms (in principle) by the motion, number, bulk, *etc.* by the action of atoms in the void.

§2. Aristotle

The classical "atomic theory" represented a simple, arguably coherent, and empirically adequate explanatory world-picture. It is conceptually satisfying, since it represents the simplest reconciliation of the facts of persistence and change—"all homogenous elements of Being are thought as unchangeable, but at the same time as in a state of motion that is self-originated."[4] At the

[3] Ibid., 90.

[4] Ibid., 91. Atomism shared the assumptions of the Eleatics and thus inherited the same problem: how could, even in principle, such a world-picture explain human sense-experience? How could atoms moving in a void account for the complex and subtle qualities we experience in perception? Democritus who was aware

same time, however, classical atomism was unabashedly rationalistic, proceeding from the phenomena to infer fanciful explanations far outstripped by the evidence available at the time. The fact that some of these early thinkers were prescient about developments in science that would occur in the third millenium after their deaths may be intriguing to us, but their speculations were nevertheless arbitrary, more the product of lucky guesses than scientific insight. By contrast, the Aristotelian corpus contains, among its philosophic riches, an analysis of causality informed by extensive empirical investigation. His inquiries into causal processes—especially into the processes of "becoming" characteristic of living things, *i.e.*, their growth and maturation, signaled a methodological reorientation toward the natural perceptible world, about which knowledge based on careful and systematic observation was possible.

The following, in no particular order, are several Aristotelian theses that express a consistently realistic approach to the ontological foundations of causality. They reflect an undercurrent of thinking—a particular conception of the natural world—which flows from an immersion in scientific investigation. Aristotle's philosophical writings often disclose a disdain for Platonic transcendentalism, which puts the validity of ontological concepts outside the reach of testing by experience. He insists, on the contrary, that our ontology, including our theory of causality, be responsive to the conceptual needs of scientific reasoning and scientific explanation. I will adopt this as a guiding theme throughout the following.

1. Intrinsic Natures as Causal Primaries. Aristotle is concerned to identify the *primary* causal principle operative in an entity of a given sort, which is that entity's "nature," (φυσις) so what a thing's nature is should be understood in various senses depending upon the kind of entity under consideration. The nature of a living thing is "the primary immanent element in a thing, from which its growth proceeds."[5] This primary immanent element is the organism's "form," exemplified in the mature phase of the development of an organism at "the end of the process of becoming."[6] The process of maturation is a process of potential for form becoming actualized. For non-natural objects, "nature is the primary matter of which [it] consists or out of which it is made."[7] In this summary of his position, Aristotle emphasizes the sense in which nature always refers to internal determinants of motion and change:

> From what has been said, then, it is plain that nature in the primary and strict sense is the substance of things which have in themselves, as such, a source of movement; for the matter is called the nature because it is qualified to receive this, and processes of becoming and growing are called nature because they are movements proceeding from this. And nature in this sense is the source of the movement of natural objects, being present in them somehow, either potentially or actually.[8]

In short, a nature is a "principle or cause of being moved and of being at rest in that to which it belong primarily, in virtue of itself and not accidentally."[9]

Following the nomenclature of *Topics*, a nature may be an essence or a property—the former

---

of the need for some sort of explanation, attempted to articulate a prototype of an early version of a physiological theory of visual perception, according to which atoms emanating from substances impacted upon our eyes, producing subsequent motions in our bodies, which were identical with the sensations experienced.

[5] Aristotle *Metaphysics* Δ.4.1014b17-18.

[6] Ibid., 1015a11; idem, *Physics* II.1.194a29.

[7] *Metaphysics* 1014b28.

[8] Ibid., 1015a13-19.

[9] *Physics* II.1.192b21-23.

if the nature is definitive of the thing, the latter if not. For Aristotle, both essences and properties are constitutive properties of a thing—those aspects of a thing that secure its being the kind of thing that it is. "Accidents," on the other hand, are not constitutive properties of things, being only "relative" or "temporary" properties. Alterations of an entity's attributes *per accidens* are consistent with the persistence of its constitutive properties, *i.e.,* with its particular nature. If the nature of a thing is that inner principle or cause of motion and superficial change, then these causally fundamental aspects of an entity's constitution determine the kinds and range of accidental or superficial changes possible for an entity.

While the phrase "essential nature" is sometimes used to contrast the causally primary properties of a thing from its derivative properties, the term welcomes conflation of ontic and epistemic issues, and is best avoided. In brief, "essence" is an epistemological concept, "nature" is a metaphysical one. Issues concerning what is or is not an essential property of a thing are definitional issues, driven by "questions of sameness and difference,"[10] and are issues of categorization or classification. What the *nature* of a thing is, on the other hand, is what provides causal explanations for a things' behavior, and may or may not be relevant to a thing's essence.

2. The Ontic Conception of Scientific Explanation. The first theory to maintain that "explanatory knowledge is knowledge of the causal mechanisms … that produce the phenomena with which we are concerned"[11] was Aristotle's:

> Scientific knowledge (*episteme*) can only be achieved by one who can successfully identify the cause (*aitia*) that makes a fact (*pragma*) be what it is; he must know the cause from which the fact results … [12]

Causality and scientific explanation remain as philosophically intertwined today as they were in antiquity. A cause is an explanatory factor—something that we "blame" the effect on. According to Aristotle, such blame may be placed upon up to four different aspects of things; "as things are called causes in many ways, it follows that there are several causes of the same thing (not merely accidentally)."[13] These four modes of causation are as follows:[14]

(i) *Efficient cause.* An efficient cause is that which explains the motion and change of things. Efficient causes are the makers of what is made and the changers of what is changed. The explanatory efficient cause is an entity in action—whether natural or human—which produces the event to be explained. Thus the efficient cause can be identified either by the entity acting, *e.g.,* a house-builder, or by the action of the entity, *e.g.,* a house-builder's *working*.

(ii) *Material cause.* A material cause is that which explains the persistence of things, and refers to the kind of substance of which things are composed, *e.g.,* bronze, silver.

(iii) *Formal cause.* A formal cause is that defining element or essence of a kind which explains why instances of the kind are known as variants of an archetype, and classified together.

---

[10] Aristotle *Topics* I.4.102a9.

[11] Wesley Salmon, "Four Decades of Scientific Explanation," in *Scientific Explanation*, Minnesota Studies in the Philosophy of Science, vol. XII, ed. Philip Kitcher and Wesley C. Salmon (Minnesota: University of Minnesota Press, 1989), 128.

[12] William A. Wallace, *Causality and Scientific Explanation,* vol. II (Ann Arbor: University of Michigan Press), 11. Cf. Aristotle *Prior Analytics* 71b8-11.

[13] *Physics* II.3.195a4-5.

[14] Cf. *Physics* II.3 and *Metaphysics* Δ.2

(iv) *Final cause*. A final cause explains the purpose or function of an activity or thing; a final cause is an "end or that for the sake of which a thing is done."

Aristotle's celebrated "bronze statue" example identifies unique causal contributions from each of the four possible types of cause: its efficient cause—the sculptor's art; its material cause—the bronze; its formal cause—a human shape; its final cause—representation of beauty. Notoriously, such clear-cut illustrations are the exception rather than the norm, being limited for the most part to artifacts. In animals, the distinction between formal and final cause collapses, while in the *Meteorology*, final causality is absent from all nonbiological terrestrial phenomena. Despite the absence of a clear account of how many types of causes are required to explain some state of affairs, it is quite clear that for any explanation, at least one type of cause must be cited. Explanation cannot avoid introducing causality.

Not all theories of scientific explanation have granted causality a focal role. In the 1960s, the philosophically most fashionable position was one that dismissed causality as irrelevant to explanation altogether, and focused on explicating the logical structure and criteria of adequacy of explanatory arguments within the framework of the deductive-nomological model.[15] The D-N theorists thought that scientific explanations were deductive arguments, but *so did Aristotle.* Nonetheless, he believed that the *central* task of explanation was getting the causes of things right. Aristotle's theory of scientific explanation has a dual focus; it is concerned with identification of causes, but it is also concerned with the deductive subsumption of particular facts by more general principles[16]—and the more universal the premises of the syllogism, the better the explanation. Scientific explanations should seek to embed the phenomenon to be explained in the most comprehensive theory possible. Aristotle's syllogistic analysis in the *Posterior Analytics* and the "four causes" theory of explanation expose different aspects of explanation, so there is no basis for assuming that these two perspectives must come into direct conflict at some point. Ultimately however, the concern with causal analysis dominates:

> The theory of the four causes is the more comprehensive theory. Aristotle maintains that the four causes are the *only* modes of explanation. (*Phys.* II 3, II 7; *Metaph.* I 10), whereas he does not maintain that all explanation has a syllogistic structure. At most, he argues in *An. Post.* II 11 that any of the four causes can feature in a syllogism as the middle term … and can explain the conclusion.[17]

For Aristotle, the perspective that informs the theory of scientific explanation is essentially the "ontic conception," to use Salmon's term.

3. Anti-Reductionism and the Ontological Unity of Organisms. Aristotle took living things as paradigmatic of particulars. It so happens, of course, that living things are structurally fairly complex—that their "material causes" are parts of the living whole. Reductive materialism focuses solely on the material composition of things, and takes what is physically irreducible—*i.e., simple*—as the essence of particularity. The unity of simple particulars requires no explanation, since they have no parts. For Aristotle, the unity of living things is explained by "formal" principles of internal organization, and their growth is coordinated by and directed toward an intrinsic *telos*—an irre-

---

[15] Causal laws could be accomodated within the D-N schema, however.

[16] These two perspectives are referred to as the "ontic" and "epistemic" conceptions of explanation in Wesley Salmon, *Scientific Explanation and the Causal Structure of the World* (Princeton: Princeton University Press, 1984).

[17] Richard Sorabji, *Necessity, Cause, and Blame: Perspectives on Aristotle's Theory* (Ithaca, NY: Cornell University Press, 1980), 56.

ducible potential for form. Form is the integrating element that differentiates individual particulars from "piles of stuff." While the discoveries of chemistry, anatomy, physiology and evolutionary biology have eliminated the need for a concept referring to the intrinsic *telos* and the ontologically distinct (albeit physically inseparable) form, the causal realist ought to be impressed by the argument that since living things are organized in a manner such that they can maintain their own structural integrity and interact causally with other entities as units, they can therefore legitimately be regarded as particulars.

4. On Natural Necessity and the Scope of Determinism. At some points, Aristotle suggests that causes must *necessitate* their effects. The realist conception of necessity in causality can be traced to Aristotle's belief that for an entity with a certain capacity (δύναμις), under specific circumstances, there is only one action possible for it. This claim needs a crucial qualification, however, for

> the non-rational potentialities are all productive of one effect each, but the rational produce contrary effects, so that they would produce contrary effects at the same time; but this is impossible. That which decides then, must be something else; I mean by this, desire or choice.[18]

The point of this section of Book Θ is to identify the different conditions under which natural necessitation occurs in persons *vs.* other living and non-living things. Since human beings may at any time have several contrary (non-compossible) "appetites" or "rational wishes," at most only one of which can be satisfied at any given point in time, the decision to act in any one way presupposes some process of deliberation. Once a choice to act has been made (assuming the person in fact has the capacity to act in such a way), and further, that the conditions for the actualization of the capacity are in place, then the occurrence of no other action but the one chosen is possible. In Aristotle's typically compact expression: "everything which has a rational potentiality, when it desires that for which it has a potentiality and in the circumstances in which it has it, must do this."[19] This passage implies a two stage action theory. In the first stage, a decision functions as a triggering cause of the actualization of one of the simultaneously present potentials for action. This stage is under volitional control. In the second stage, the same decision initiates a causal process whose outcome is necessitated by the antecedent decision as well as by causally simultaneous factors. The view of persons most in accordance with this model of action regards them as agents who have free volitional *control* over their actions *by means of* a deterministic causal connection between a decision to act and the bodily motion that follows upon it. For non-rational (and therefore non-volitional) creatures and other inanimate objects, the nature of their agency is qualitatively different: they are wholly deterministic.

The division of things into those that have rational and non-rational δύναμις implies a corresponding modal distinction between facts and events that are contingent *vs.* those that are necessary—*i.e.*, between events and facts whose existence are necessitated by the nature of things, and which could *not* have been otherwise than they are, and those whose existence depends upon the outcomes of human deliberations and actions which *could* have been otherwise than they are. This shows Aristotle's concern for the "modal" aspects of scientific explanation as well, and confirms Salmon's suggestion that we can "distinguish three basic conceptions of scientific explanation—modal, epistemic, and ontic—that could be discerned in Aristotle, and that have persisted

---

[18] *Metaphysics* Θ.5.1048a8-10.

[19] Ibid., 1048a13-15.

down through the ages."[20]

5. Capacities and Liabilities. At *Metaphysics* Δ.12, an account of the term δύναμις or "power"[21] is again presented, in terms similar to "nature." While a nature is a source of change or movement in a thing in virtue of itself, δύναμις is "a source of change or movement in another thing or in the same thing *qua* other, and also the source of a thing's being moved by another thing or by itself *qua* other."[22] The distinction between natures and powers is a distinction between intrinsic properties and relational properties of things. Powers come in two sorts, depending on whether the capacity is performative or passive. Aristotle's treatment of passive powers as the absence of capacity is surprisingly nuanced. While things that resist change are such in virtue of their capacities, "things are broken and crushed and bent and in general destroyed not by having a capacity but by not having one and by lacking something."[23] Things that have a disposition to suffer alteration by the actions of others things fail to have a capacity to resist the action of the things that affect them. In this sense, Aristotle describes a liability as a "privation," or the absence of a positive principle. Nonetheless, it is legitimate to say of some particular that is *has* such a privation, as this example illustrates:

> Even that which perishes is thought to be capable of perishing, for it would not have perished if it has not been capable of it; but, as a matter of fact, it has a certain disposition and cause and principle which fits it to suffer this.[24]

This implies that there is some "cause and principle," *i.e.*, some specific aspects of a thing's nature in virtue of which both capacities and liabilities exist in things.

The Aristotelian theory of causality was influential among medieval philosophers notably Aquinas, and Averroes.[25] Kogan summarizes Averroes' theory in this way:

> … things both persist and change in the way they do, because they have essential natures which are properly their own. Under analysis, these natures emerged as hierarchies of active and passive powers or what we may now call simply powers and dispositions. It is the distinctive selection and arrangement of these powers, included in the structure of an entity, that allows us to explain why different individuals belong to the same or different natural kinds. … Since the theory holds that these structures express the possible combinations of powers and dispositions *in rerum natura*, it follows that they also set continuing limits upon both the kinds of effects particulars may produce and the changes they can undergo.[26]

I contend that this picture of powers and natures is basically correct, as far as it goes, and that it provides the rudimentary conceptual basis upon which a realist theory of causality ought to be formulated.

---

[20] Salmon, "Four Decades," 121.

[21] The best translation for this term is "power." "Potency" and "potentiality" have a slightly different sense, since they imply a correlative absence of actuality. Likewise, act implies an absence of potency.

[22] *Metaphysics* 1019a18-20.

[23] Ibid., 1019a27-8.

[24] Ibid., 1019b3-5.

[25] Aristotle's influence on the medievals is discussed in Wallace, *Causality and Scientific Explanation*, 6; Rom Harré and Edward H. Madden, *Causal Powers* (Totawa, NJ: Rowman and Littlefield, 1975), 98-100); Barry S. Kogan, *Averroes and the Metaphysics of Causation* (New York: SUNY Press, 1985).

[26] Kogan, *Averroes*, 166.

§3. Locke

This Aristotelian conception of the natures of things has not always been uncontroversial. If, as some seventeenth-century critics of Aristotelianism[27] believed, in order for an explanation to be scientific, it must make essential reference to the natures of things, and those natures themselves are not open to observation, then scientific explanations advert to hidden, unobservable structures and principles that are closed to empirical investigation. These "natures" would be, in essence, "occult" qualities, and science would have nothing to say about them. Their existence and attributes could not be settled by reference to observable fact. What the seventeenth-century critics of scholasticism may have objected to as "occult" in Aristotle's metaphysics was the idea of a potential for form, understood as an immaterial principle of actualization or "becoming." Whatever the details of the source of their discomfort, let us assume that the critics of Aristotelian science were on target insofar as they insisted that only the discoverable qualities of actual, sensible matter (and its parts) are admissible in scientific explanations.

It is somewhat ironic, therefore, that John Locke, one of the leading seventeenth century defenders of Enlightenment empiricism, developed a notion of causal power that bears the ancestry of the medieval Aristotelians. Locke begins with the assumption that bodies produce ideas in us "manifestly by impulse, the only way which we can conceive bodies to operate in."[28] In other words, only a transfer of motion—*i.e.,* momentum—could explain how a body might affect another. This served to restrict, by philosophical stipulation, the range of tenable analyses of causation, and with it, the criteria of explanatory adequacy. Within this framework, Locke reasoned that since the objects of sensory awareness are not in any form of immediate connection to the mind and yet we become aware of them in specific sensory forms, it is the case that

> some motion must be thence continued by our nerves, or animal spirits, by some parts of our bodies, to the brain or the seat of sensation, thereafter to produce in our minds the particular ideas we have of them. And since the extension, figure, number, and motion of bodies of an observable bigness, may be perceived at a distance by the sight, it is evident some singly imperceptible bodies must come from them to the eyes, and thereby convey to the brain some motion which produces these ideas which we have of them in us.[29]

This argument for the corpuscular hypothesis is a form of inference to the best explanation. Locke attempted to present an explanation of *how* knowledge of the properties of bodies is possible. Given that such knowledge *is* possible, if bodies can only interact causally with other bodies of approximately the same mass and extension, and if the only "conceivable" way for bodies to transfer motion to other bodies is via mechanical impact, then all bodies must be composed of a multitude of imperceptibly small corpuscles.

The analysis of the transfer of motion or momentum from one body to another was, at the time, fairly well understood. Only a small number of parameters were needed to model mechanical interactions. These so-called "primary qualities" were supposed to be able to explain both the sensible qualities of bodies such as their colour, taste, *etc.,* as well as their capacities to affect others

---

[27] Pierre Gassendi and Joseph Glanvill's anti-Aristotelianism in particular is noted in Nancy Cartwright, "Aristotelian Natures and the Modern Experimental Method," in *Inference and Method,* ed. John Earman (Pittsburgh: University of Pittsburgh Press, 1993), 45.

[28] John Locke, *An Essay Concerning Human Understanding,* vol. 35 of *The Great Books of the Western World,* ed. R. M. Hutchins (Chicago: Encyclopedia Brittanica, 1952), 135.

[29] Ibid., 135.

in determinate, consistent ways, such as the power of opium to induce sleep. Primary qualities are five in number: extension, solidity, figure, motion, and number. They are *essential* properties of material bodies in the sense that we cannot conceive of bodies existing without each of these qualities in some measure, whether these bodies are sensible or not. They are also *intrinsic* properties, in that they "really exist in those bodies" whether "any one's senses perceive them or no."[30] Finally, the primary qualities of bodies generate accurate representations of themselves in the mind through their ability to structure circumambient corpuscles of matter and impel them toward the senses in the right way. The resemblance of the primary qualities to the associated visual appearances (ideas) that the primary qualities produce is the ground for whatever claims to epistemic objectivity we might be able to make for our knowledge of things.

Locke divided the qualities of bodies into three types, of which the primary qualities are, of course, the first. The second and third types of qualities are the two kinds of "powers" which are determined by, and realized in, the primary qualities of objects' microscopic parts. The first kind, dubbed "secondary qualities," are those aspects of objects which are "nothing in the objects themselves" but the powers "to produce various sensations in us by their primary qualities, *i.e.*, by the bulk, figure, texture, and motion of their insensible parts, as colours, sounds, tastes, etc."[31] Implicitly, Locke thereby affirms that powers are *causal* powers. The second category of powers is comprised of those that produce effects in everything *but* the sensory organs. These "tertiary" qualities are nothing but the powers of objects, in virtue of the primary qualities of their insensible parts, to affect other objects after the manner of the mechanical interaction of two colliding macroscopic objects.

Locke does not *identify* the powers of bodies with the particular extension, figure, bulk, number and motion of the insensible parts of things. Rather, the latter *explain* the former. The primary qualities of the insensible parts of a thing determine why the thing—the body—has the effects that it does on other objects. The concept of "power" denotes the capacity or liability that a body has to act in certain specific ways in specific situations because its microscopic parts have the primary qualities that they do. There is no separate entity called a "power" above and beyond the constituent parts of bodies; there is no power *in* a body, rather, there are powerful bodies.[32] This is consistent with Locke's theory, expressed in a separate section of the *Essay*, of the origin of our idea of power:

> The mind being every day informed, by the senses, of the alteration of those simple ideas it observes in things without; and taking notice how one comes to an end, and ceases to be, and another begins to exist which was not before; … and concluding from what it has so constantly observed to have been, that the like changes will for the future be made in the same things, by like agents, and by the like ways,—considers in one thing the possibility of having any of its simple ideas changed, and in another the possibility of making that change; and so comes by the idea which we call *power*.[33]

The idea here is that the concept of power refers to whatever internal factor is responsible for the possibility of the production of changes in the observable features ("the ideas we have") of things. Power, "thus considered, is two-fold, viz., as able to make, or able to receive any change. The one

---

[30] Ibid., 135.

[31] Ibid., 134.

[32] James Humber and Edward Madden, "Natural Necessity," *New Scholasticism* 47 (1976): 214-217.

[33] Locke, *Essay*, 178.

may be called *active*, and the other *passive* power." This distinction between active and passive powers corresponds to what I have been referring to as capacities and liabilities.

Despite our need for a concept of power to refer to the underlying explanatory factors determining the way things act, Locke is quite pessimistic about the prospects for the scientific study of powers.[34] Macroscopic bodies are observable because their individual primary qualities produce faithful representations of themselves in our minds. Conversely, we can acquire knowledge of the primary qualities of things based on the ideas we have of them. Unfortunately, we cannot acquire specific knowledge of the powers (secondary and tertiary qualities) of things, because the relevant primary qualities are of particles too minute to be able to deliver faithful representations of themselves to us. Since the basis of powers resides in a body's minute parts, we cannot know exactly what specific shape, motion, *etc.* of its corpuscles is responsible for generating specific ideas, *e.g.*, the colour violet or the taste of sweetness. If we could somehow "see" the bulk, figure, motion, *etc.* of the insensible particles of which things are made, then that would satisfy Locke's condition for empirical knowledge of their powers:

> Did we know the mechanical affections of the particles of rhubarb, hemlock, opium, and a man, as a watchmaker does those of a watch, whereby it performs its operations … we should be able to tell beforehand that rhubarb will purge, hemlock kill, and opium make a man sleep.[35]

Owing to the minuteness of corpuscles, we establish neither any specific connections between specific corpuscular configurations and the ideas they presumably induce in us, nor the changes they are capable of producing in other things. All that we are justified in claiming is that there is *some* such connection for each power. "For, if sugar produce in us the ideas which we call whiteness and sweetness, we are sure there is a power in sugar to produce those ideas in our minds, or else they could not have been produced by it."[36] While the existence of powers is never in question, the real constitution of the minute parts of bodies on which powers depend is inscrutable. There is therefore a sort of empirical vacuity in Locke's notion of "power," since it refers to qualities which have no necessary connection to any observable consequences.[37]

The Lockean "complex idea" of a substance is the sum of simple ideas of qualities found to regularly co-exist in specimens of the substance. Consider the substance we call "gold" (one of Locke's favorite examples). While we are naturally inclined to think that gold is gold because of its possession of some unique inner structure, that inner structure would be undiscoverable. We call certain specimens of gold "gold" not on the basis of knowing their microstructure, but because they share in a pattern of co-instantiation of specific observable properties: yellowness, malleability, high density, *etc.* We cannot know whether gold's characteristic observable properties are anything more than a contingent co-instantiation of property tokens, because of our incurable ignorance concerning the relationships of the observed qualities of the metal to the specific "affections" of gold's insensible parts that give rise to, in us, the several ideas we receive upon acquaintance with it.

> For of all the qualities that are co-existent in any subject, without this dependence and evident connexion of their ideas one with another, we cannot know certainly any two

---

[34] See Ernan McMullin, "Structural Explanation," *American Philosophical Quarterly* 15, no. 2 (1978): 139-147 for a poignant contrast of Locke's pessimism with Newton's optimism.

[35] Locke, *Essay*, 321.

[36] Ibid., 239.

[37] Locke's own idiom labels such ideas as "inadequate." Inadequate ideas are those that are only "partial or incomplete representations of those archetypes to which they are referred."

> to co-exist, any further than experience, by our senses, informs us. Thus, though we see the yellow colour, and, upon trial, find the weight, malleableness, fusibility, and fixedness that are united in a piece of gold, yet; because no one of these ideas has any evident dependence or necessary connexion with the other, we cannot certainly know that where any four of these are, the fifth will be there also, how highly probable soever it may be; because the highest probability amounts not to certainty, without which there can be no true knowledge.[38]

What is it that explains the regular co-occurrence of qualities in gold, such as the shiny yellow colour, the high density, and its liability to discoloration by mercury? People presume, Locke reports, that this phenomena of co-existence of qualities is the effect in each specimen of a type of substance having a unique real essence, from which those qualities "flow." The real essence of a substance is its "internal constitution" upon which the characteristic sensible qualities of specimens—the "nominal essence"—of that substance depend. The real essence of a substance or thing refers to the common *nature* had by each token of the substance type—the specific "figure, size, and connexion of its solid parts" that realizes the power of the particulars to produce just those co-occurring sensible qualities by which we type-identify the substances that exemplify them.

Aristotle bifurcated the causal determinants of the potentialities of things into matter and form, with the intrinsic form constituting an ontologically independent explanatory factor possessing the ability to mould matter in certain ways. The concept of form was an epistemic convenience, because it provided an objective basis for the classification of species into natural kinds. Locke regards his theory of real essences as an advance on the Aristotelian theory of substance for two reasons. First, it complies with the general empiricist constraint that abstract general ideas must be reducible to perceptible aspects of reality:

> If any one will say, that the real essence and internal constitution on which these [secondary] properties depend, is not the figure, size, and arrangement or connexion of its solid parts, but something else, called its particular *form*, I am further from having any idea of its real essence than I was before. For I have an idea of figure, size and situation of solid parts in general, … but when I am told that something besides the figure, size, and position of the solid parts of that body in its essence, something called *substantial form*, of that I confess I have no idea at all.[39]

Second, it avoids the empirical embarrassments to Aristotle's theory of substance which Locke describes here:

> The frequent productions of monsters, in all the species of animals, and of changelings, and other strange issues of human birth, carry with them difficulties, not possible to consist with this hypothesis; since it is as impossible that two things partaking exactly of the same real essence should have different properties, as that two figures partaking of the same real essence of a circle should have different properties.[40]

Locke's concept of "real essence" represents a reintegration of form and matter as causes of the powers of particulars, yielding a more robust conception of the ontological unity of substance, and thereby confirming our assumption that the compositional and structural aspects of a substance's nature are separable only by abstraction.

Locke regarded the inscrutability of powers as a permanent condition. He thought that their unknowability was a consequence of the inherent limitations of the senses and reason. Fortunately

---

[38] Locke, *Essay*, 316.

[39] Ibid., 241.

[40] Ibid., 259.

for science, Locke's pessimism was unwarranted. The boundary between what is observable and what is not is a function primarily of the state of technology, and secondarily of experimental sophistication. As our tools become smaller and more precise, our measuring devices become more sensitive and accurate, our microscopes more powerful, ever more delicate and subtle structural features of reality are being brought into view.

In the face of scientific progress, drawing the distinction between real and nominal essence on the basis of a fundamental distinction between the observable and unobservable "parts" of things is untenable. On the other hand, Locke was right to draw a distinction between the intrinsic constitution of things (their natures) and the irreducibly relational phenomena in which those things are involved (the actualizations of their powers), where the latter are dependent upon the particulars' intrinsic constitutions. This distinction is key to the realist ontology.

1.2 Hume's Account of Causation

J. L. Mackie[41] presents an overview of Hume's argument for the ultimate conclusion that causation "in the objects" is nothing but the regular succession of events. Mackie embarks on the ambitious and laudable task of diagramming the argument in which Hume's analysis of causation is given, specifying the key premises, various levels of sub-conclusions, and inferential moves that Hume makes. Mackie identifies the following passage from *A Treatise of Human Nature* as providing a capsule summary of the main threads of Hume's argument:

> Before we are reconciled to this doctrine, how often must we repeat to ourselves, *that* the simple view of any two objects or actions, however related, can never give us any idea of power, or of a connexion betwixt them: *that* this ideas arises from the repetition of their union: *that* the repetition neither discovers nor causes anything in the objects, but has an influence only on the mind by that customary transition it produces: *that* this customary transition is therefore the same with the power and necessity; which are consequently qualities of perceptions, not of objects, and are internally felt by the soul, and not perceived externally in bodies?[42]

Mackie takes the statements in this passage to constitute the "backbone" of the whole overall structure of Hume's account, and develops the supporting sub-arguments for each of the premises. The presence, in the *Abstract*, of a symmetrical skeletal presentation lends additional support to the propriety of this approach, Mackie claims.

My own reconstruction of Hume's argument differs from Mackie's slightly, relying primarily on the argument of §§ IV, V, and VII of *An Enquiry Concerning Human Understanding*, which Hume himself claimed superseded Book I of the *Treatise*.[43] The overall argument has three main threads, the conclusions of which appear below as (**A**), (**B**) and (**E**). These are the argument from the secrecy of causal powers, the argument from inductive skepticism, and the argument from the lack of empirical content of the notion of necessary connection. I will first state the central argument and

---

[41] J. L. Mackie, *The Cement of the Universe: A Study of Causation* (Oxford: Clarendon Press, 1980), ch. 1.

[42] David Hume, *A Treatise of Human Nature*, 2d ed., ed. L. A. Selby-Bigge, rev. P. H. Nidditch (Oxford: Clarendon Press, 1978), 166.

[43] Philosophers generally ignore Hume's own repudiation of the *Treatise* as a "juvenile" and unseasoned work, and treat both it and the *Enquiry* as roughly equivalent expressions of a single philosophical doctrine. If a preference for one over the other is present in philosophical circles, it is for the more detailed early work. Hume likely did exaggerate the substantive differences between Book I of the *Treatise* and the *Enquiry*, but the differences that *are* present have subtle substantive implications. Given Hume's own express preference for the content and style of the *Enquiry*, scholars ought to give pride of place to the later work.

discuss the premises themselves, just briefly enough to expose the areas where Hume is particularly vulnerable to confutation.

§1. The Structure of the Argument

The conclusion of the argument is expressed succinctly in this passage:

> Our idea, therefore, of necessity and causation arises entirely from the uniformity observable in the operations of nature, where similar objects are constantly conjoined together, and the mind is determined by custom to infer the one from the appearance of the other.[44]

The thread of the argument of *Enquiry IV* is this:

(1) Events in nature ultimately are produced by the ("secret") causal powers of things.

(2) No analysis of the sensible qualities of objects will tell us about their powers.

(3) It is possible that the sensible qualities of objects may remain constant while the secret powers with which they have been associated may change.

∴ (**A**) There is no known connection between the sensible qualities and secret powers of objects.

The second thread of argument is as follows:

(4) Causes and effects are distinct events.

(5) Any number of effects can be conceived to follow a prior event.

(6) A change in the course of nature is conceivable in the imagination.

∴ (**B**) The future may not resemble the past.

(**A**) and (**B**) together imply:

(**C**) We cannot logically infer effects from observable causes.

The part of *Enquiry V* needed for the argument can be stated very quickly:

(7) Beliefs are lively ideas, to which the mind moves easily as a result of a present impression with which the idea is associated.

∴ (**D**) We have a habit of belief by which we anticipate events on the basis of observing the antecedents with which they are constantly conjoined.

In *Enquiry VII* Hume presents the following:

(8) The validity of any idea can be established by the decomposition of the idea into its simple components and determining the impression(s) of which the ideas are copies.

(9a) While considering a sequence of two events, we are never able to apprehend any powers possessed by the cause in virtue of which the effect necessarily follows.

(9b) While considering the actions which issue from our volitions, we are unaware, in consciousness, of any power of the mind that binds to our volitions the actions that we notice regularly follow from them.

---

[44] David Hume, *An Enquiry Concerning Human Understanding*, ed. Eric Steinberg (Indianapolis: Hackett Publishing Co., 1977), 55.

(9c) While considering the succession of ideas in a process of conscious deliberation, we are unaware of the power of the mind which binds our thoughts together in a way that makes an idea the necessary consequence of a prior one.

∴ (9) No impression of causal power or necessary connection is discoverable in experience.

(**8**) and (**9**), together with (**D**) imply:

(**E**) The impression from which our ideas of causation and necessity originate are in the mind, (not the objects) as felt habitual transitions from observed causes to anticipated effects in the imagination.

After having extracted the main lines of supporting argument, the "capsule summary" of the main argument can be presented:

(**C**) IF we cannot logically infer effects from observable causes; and

(**D**) IF we nonetheless have a habit of belief by which we anticipate events on the basis of observing the antecedents with which they are constantly conjoined; and

(**E**) IF the impression from which our ideas of causation and necessity originate are in the mind, (not the objects) as felt habitual transitions from observed causes to anticipated effects in the imagination;

THEN, it follows that causation in the objects is, so far as we know, nothing but the constant conjunction of events.

§2. Explication of Premises

(1) Hume has often been interpreted as arguing for the conclusion that causality (in the objects) is just constant conjunction. But that directly contradicts the numerous places in the *Treatise* and the *Enquiry* where Hume clearly maintains that causality in the objects is what *accounts for* the existence of the constant conjunctions that we observe. Hume is making the epistemological point that our powers of understanding are insufficient to penetrate into the inner structure of things to expose to the senses the "secret springs and principles" of their operation. Nevertheless, the metaphysical point that causality in the objects *is* the operation of intrinsic necessitating causal powers is literally overwhelming.[45] Here are just a few illustrative references:

> Nature … conceals from us those powers and principles on which the influence of … objects entirely depends.[46]
>
> Is there not … either in a spiritual or material substance, or both, some secret mechanism or structure of parts, upon which the effect depends …?[47]
>
> … my intention never was to penetrate into the nature of bodies, or explain the secret causes of their operations.[48]
>
> … reason fails us in the discovery of the ultimate connexion of causes and effects …[49]

---

[45] This thesis is convincingly argued in Galen Strawson, *The Secret Connexion: Causation, Realism and David Hume* (Oxford: Oxford University Press, 1989).

[46] Hume, *Enquiry*, 21.

[47] Ibid., 45.

[48] Hume, *Treatise*, 64.

[49] Ibid., 91.

> We have no other notion of cause and effect, but that of certain objects, which have been always conjoin'd together …. We cannot penetrate into the reason of the conjunction.[50]
>
> … in no single instance the ultimate connexion of any objects is discoverable, either by our senses or reason, and that we can never penetrate so far into the essence and construction of bodies, as to perceive the principle on which their mutual influence depends.[51]
>
> And experience only teaches us, how one event constantly follows another; without instructing us in the secret connexion, which binds them together, and renders them inseparable.[52]
>
> We are ignorant of those powers and forces, on which this regular course and succession of objects totally depends.[53]
>
> The scenes of the universe are continually shifting, and one object follows another in an uninterrupted succession; but the power or force, which actuates the whole machine, is entirely concealed from us, and never discovers itself in any of the sensible qualities of body.[54]

Galen Strawson notes, "it would be very odd if [Hume] also thought that the 'power or force' which on his view 'actuates the whole machine … of the universe' did not exist."[55] What is even odder is this: if Hume believed that the secret powers and "connexions" were totally shut off from experience and completely inscrutable, how does he know that they exist? The answer lies in the subtitle of Hume's *A Treatise of Human Nature:* "Being an Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects." By the "experimental method of reasoning," Hume has in mind the methods practiced by Newton.[56] Hume is unremittingly cautious about the specifics of the ontology presupposed by Newtonian scientific method and reasoning. Nonetheless, Hume is compelled, by a higher-order methodological premise, to accept the existence of those aspects of the world whose existence and nature are unfathomable on narrower empiricist grounds. The interesting question, I think, is whether the tension between Hume's realist methodological presuppositions and his empiricist epistemological aspirations is tolerable, or whether it reveals a deeper intolerable incoherence in his overall account.

(2) Hume's point in this premise is to show that the powers of a thing cannot be deduced, *a priori*, from a specification of the sensible qualities of a thing. As an example, Hume asks us to consider that "[o]ur senses inform us of the color, weight, and consistency of bread, but nei-

---

[50] Ibid., 93.

[51] Ibid., 400.

[52] Hume, *Enquiry*, 43.

[53] Ibid., 37.

[54] Ibid., 42.

[55] Strawson, Secret Connexion, 14.

[56] See Hume, *Treatise*, 639; also Tom L. Beauchamp and Alexander Rosenberg, *Hume and the Problem of Causation* (Oxford: Oxford University Press, 1981), 43.

ther sense nor reason can ever inform us of those qualities which fit it for the nourishment and support of the human body."[57]

(3) While deduction is incapable of demonstrating that an object with certain sensible qualities will be associated with specific powers, neither can induction. Hume writes,

> Let the course of things be allowed hitherto ever so regular, that alone, without some new argument or inference, proves not that for the future it will continue so. In vain do you pretend to have learned the nature of bodies from your past experience. Their secret nature and consequently all their effects and influence, may change without any change in their sensible qualities.[58]

Here is Hume's example:

> The bread which I formerly ate nourished me; that is, a body of such sensible qualities was, at that time, endued with such secret powers. But does it follow that other bread must also nourish me at another time, and that like sensible qualities must always be attended with like secret powers? The consequence seems nowise necessary.[59]

Hume's suggestion here is that it is conceivable that the secret powers of a thing could undergo considerable alteration, while correlative differences in sensible qualities could go undiscerned. Another example might be of a tablet containing a small amount of a drug that breaks down and loses its effectiveness in a manner of weeks. A growing but undetectable microfracture within a metal support beam could jeopardize the structural soundness of a building. On the other hand, the quotations from the section detailing (1) above would seem to imply that the powers of a thing cannot change without a correlative change in the inner constitution of a thing which determines its causal powers, even if at the level of unaided observation everything looks the same.

(4) According to Hume, "every effect is a distinct event from its cause."[60] "All events seem entirely loose and separate."[61] Hume's thesis here is essentially a psychological claim about the relations between events as perceived by us. Perceptible events are impressions, and as such, they are experienced as discrete, transient phenomena that become connected together and acquire temporal continuity in the faculty of imagination.[62] Insofar as this is an empirical claim about the processes of the Human mind, it is falsifiable. If it turns out that there exist physiological processes of perceptual integration in virtue of which a sequence of separate events appears to us as parts of a continuous causal process, then Hume is wrong on this point.

(5) This premise (and the next) express Hume's imaginability criterion of physical possibility. Hume asks us to consider, as an example of the principle expressed in the premise, a billiard ball moving in a straight line toward another:

---

[57] Hume, *Enquiry*, 21.

[58] Ibid., 24.

[59] Ibid., 21.

[60] Ibid., 19.

[61] Ibid., 49.

[62] This is explored in more detail in my critical remarks concerning Hume's sensationalism in §1.3.1.

> [A]s the result of their contact or impulse, may I not conceive that a hundred different events might as well follow from that cause? … All these suppositions are consistent and conceivable. Why then, should we give the preference to one which is no more consistent or conceivable than the rest?[63]

Any judgment that is partial to one conceivable possible outcome is arbitrary. In the *Treatise*, Hume is more blunt: "Any thing *may* produce any thing."[64] Hume is attempting to establish, with this sub-argument, that one cannot deduce, on the basis of the observable properties of things, what a thing *must* do. There is no necessary connection between the nature of a thing and its powers.

(6) Just as we cannot deductively establish what things must or must not do, we cannot establish it inductively either. Stability or uniformity of behavior is insufficient, by itself, to rule out the possibility of the occurrence of the unexpected, let alone the bizarre. As Hume argues: "May I not clearly and distinctly conceive that a body, falling from the clouds and which in all other respects resembles snow, has yet the taste of salt or feeling of fire?" He continues: "whatever is intelligible and can be distinctly conceived implies no contradiction and can never be proved false by any demonstrative argument or abstract reasoning *a priori*."[65]

(7) Section *V* of the *Enquiry*, where this premise resides, offers an analysis of our inclination to draw rationally unwarranted conclusions on the basis of experience, and concludes with a hypothesis designed to explain it. Hume offers a theory of empirical belief-acquisition, the following statement of which serves as a premise in another inferential step towards Hume's ultimate conclusion:

> Beliefs are lively ideas, to which the mind moves easily as a result of a present impression with which the idea is associated.

Hume asks:

> … what is there in this whole matter to cause such a strong conception except only a present object and a customary transition to the idea of another object which we have been accustomed to conjoin with the former? This is the whole operation of the mind in all our conclusions concerning matter of fact and existence.[66]

The application of this theory to the experience of constant conjunction leads to the hypothesis expressed in premise (**D**). This hypothesis explains "why we draw from a thousand instances an inference which we are not able to draw from one instance that is in no respect different from them"[67]—we are determined by custom or habit to do so. We acquire this habit by observing events of one kind being constantly followed by the same kind of event. This way of drawing conclusions by custom is not rationally warranted—the inference drawn is not logically valid—but is a natural disposition of the mind.

---

[63] Hume, *Enquiry*, 18-9.

[64] Hume, *Treatise*, 173.

[65] Hume, *Enquiry*, 22.

[66] Ibid., 36.

[67] Ibid., 28.

(8) Hume's impressions-and-ideas psychology, presented in §II of the *Enquiry*, plays a key role in the argument for the invalidity of the power account of causation, and for the thesis that the idea of necessity is based on an inner impression of the mind's habits of imagination. This premise states the normative implications of Hume's theory of idea-formation, in the form of an empiricist criterion of cognitive significance. This criterion is then brought to bear on the question of how our ideas of power and necessity originate.

(9) This premise is the multiply-supported subconclusion of premises (9a)-(9c), that is, that no impression of causal power or necessary connection is discoverable in experience. There is no "third thing" ever observed besides the two conjoined events—the cause and the effect. Hume surveys a number of different phenomena for the source in experience of our ideas of power and necessary connection, but to no avail.

In Hume's words,

> We have sought in vain for an idea of power or necessary connexion in all the sources from which we could suppose it to be derived. It appears that in single instances of the operation of bodies we never can, by our utmost scrutiny, discover anything but one event following another, without being able to comprehend any force or power by which the cause operates or any connection between it and its supposed effect.[68]

Premise (**E**) is itself implied by premises (8) and (9) together with (**D**). Here is premise (**E**) in Hume's own words:

> … [T]his idea of necessary connection among events arises from a number of similar instances which occur … . This connection, therefore, which we *feel* in the mind, this customary transition of the imagination from one object to its usual attendant, is the sentiment or impression from which we form the idea of power or necessary connection.[69]

1.3 Cracks in the Foundation: Two Philosophical Mistakes

Hume's positive theory of the ontology of causality is exposed to criticism on a few fronts, to which we will shortly turn. Since my primary concern is the elaboration and defense of an alternative theory of causality, I do not dwell on these criticisms. I am primarily interested in presenting the philosophical case for abandoning the Humean perspective on issues relating to causality in favour of one that is more philosophically fruitful, and more in line with the ontological presuppositions of modern scientific practice. There are two premises in Hume's case for causal anti-realism that warrant somewhat closer examination, for a couple of reasons. First, given their position within the overall logical structure of Hume's argument, the plausibility of his conclusions depends more on their truth than on other premises. Second, these key premises are, in the final analysis, not supported by sufficient argument. The premises are (i) the sensationalistic account of the origin of "ideas" and (ii) the imaginability criterion of possibility.

Showing why Hume's premises are untenable does not demonstrate the outright falsehood of his conclusions, only that his conclusions are inadequately supported by them. Nonetheless, if there are other causal anti-realist arguments to be had, it seems unlikely that they would depend upon premises which are non-Humean in spirit. This suggests that we need to investigate other accounts

---

[68] Ibid., 49.

[69] Ibid., 50.

of the origins of ideas in general (and of the idea of "possibility," in particular) in addition to alternative non-Humean theories of causality. In subsequent chapters, I present what I think is the current best case for a robustly realist theory of causality. As for the other issues, I can do little more at present than indicate which alternatives seem to be the most promising, deferring the systematic articulation and defense that they require until another time.

§1. Sensationalism as an Account of the Genesis of 'Ideas'

"Sensationalism" is the idea that the basic form of conscious awareness is a simple irreducible sensory quality, such as a colour or smell, to which all other higher forms of awareness are reducible. In Hume's work, the individual sensations are called "impressions," of which any "idea" is a copy. The idea corresponding to a simple impression is a "faint image"[70] of it which comes to mind when I see something that has the sensory quality I associate with the idea, or when I recall the term used to denote it. The complex idea is the concatenation in the imagination of simple ideas "annexed" to a particular term. According to Hume, our complex ideas about things are constructed representations ("copies") of the immediate sense-data presented to us as impressions. Since ideas are just copies of impressions, but with less "force and vivacity," they differ in their distinctness but *not* in their determinacy. This theory is a thread that runs through much of Hume's thought, and lends key argumentative support for two of the premises leading to his constant conjunction account of causality.

First, Hume's sensationalism justifies his empiricist criterion of meaning, which requires that for an idea to be meaningful it must have its origin in, and be reducible to, a simple impression or a conjunction of simple impressions. The criterion is deployed cleverly in *Enquiry VII* to support Hume's thesis of the undetectability of causality in experience. There, Hume shows that there is no impression of sensation corresponding to the ideas of "power" or "necessary connection"—no "third thing" besides the events in our outer experience from which the ideas could come. This is an essential link in the argument for the Regularity theory and the primacy of causal laws. As Beauchamp and Rosenberg maintain,

> For the purposes of Hume's argument that particular causal sequences cannot directly and immediately be recognized, all that is required is the admission that "efficacy," or power or agency, or productive force, or any of the cognates of causation, is not related to sensation in the way "red" is, whatever that may be.[71]

Hume supposed that for any idea to be meaningful, it has to be reducible to observation. That is, in principle, we ought to be able to identify sensory conditions in which it is possible to acquire the impression from which the idea derives. To be observable, an object has to have sensible qualities. If a causal connection were to be detected in experience, it would have to be a third thing "between" the two events—an existent whose properties were sensible.

Second, sensationalism supports his view that events are discrete and separate sensory phenomena, never the continuous transformation of a persistent, extended spatiotemporal entity. Therefore, we can never directly perceive powerful particulars acting in accordance with their natures.

According to sensationalism, a simple impression is an irreducible sensory quality, such as colour or smell, or a feeling. Its corresponding simple idea is a faint "image," which is recalled to

---

[70] Hume, *Treatise*, 25.

[71] Beauchamp and Rosenberg, *Hume*, 82.

the imagination under certain conditions. A complex impression is a concatenation of particular impressions of determinate sensory qualities that are united in the imagination into the ideas of objects.

> The table before me is alone sufficient by its view to give me the ideas of extension. This idea, then, is borrowed from, and represents some impression, which this moment appears to the senses. But my senses convey to me only the impression of colored points …[72]

Ideas of objects such as tables are representations, *constructed* in the imagination, which are formed by combining the ideas resulting from the impressions of spatially contiguous yet discrete coloured points. More generally, Hume declares that "our ideas of bodies are nothing but collections form'd by the mind of the ideas of the several distinct sensible qualities, of which objects are compos'd, and which we find to have a constant union with each other."[73] Hume acknowledges that while we *regard* bundles of conjoined sensible qualities as constituting objects, that appearance of unity is a result of post-impression activity of the imagination which effortlessly "unites the object within itself" so that "the fancy feels not the transition in passing from one part to another."[74]

Just as our awareness is spatially discontinuous, it is temporally discontinuous. Impressions, says Hume, are "perishing" existences[75]—they are momentary, fleeting, sensory experiences. Each moment in time "as it succeeds another [is] perfectly single and indivisible."[76] Our ideas of time and motion are derived from the manner in which the impressions appear to the mind: "time, which since it appears not as any primary distinct impression, can plainly be nothing but different ideas, or impressions, or objects disposed in a certain manner, that is, succeeding each other."[77] Thus, the motion of an object is not something of which we could have an impression. What appears to us as motion is in fact a succession of impressions of stationary objects, each one of which bears a distinguishable change in sensible qualities from the prior impression in the sequence—a change in spatial location. According to Hume,

> … as the ideas of the several distinct *successive* qualities of objects are united together by a very close relation, the mind, in looking along the succession, must be carry'd from one part of it to another by an easy transition, and will no more perceive the change, than if it contemplated the same unchangeable objects …; hence it proceeds, that any such succession of related qualities is readily consider'd as one continued object, existing without any variation.[78]

We do not *perceive* motion as the smooth, continuous displacement of a single object relative to a stationary background. To believe otherwise is to fall prey to the deception of which the vulgar are commonly victim. The proper account of our perceptual experience, once the effects of fancy have been subtracted, holds that the successive states of an object are perceived as a discontinuous sequence of impressions. Both the temporal continuity of motion and change and the spatial unity of substances are products of the natural synthetic activity of the imagination and are not simply

---

[72] This is a succinct statement of sensory atomism. (See Hume, *Treatise*, 34).

[73] Hume, *Treatise*, 219.

[74] Ibid., 221.

[75] Ibid., 194.

[76] Ibid., 34.

[77] Ibid., 37.

[78] Ibid., 220.

given in experience.

While we perceive events as disjoint, independent and discrete appearances, never as causally connected, ideas may be so connected: "ideas are associated by resemblance, contiguity and causation; and impressions only by resemblance."[79]

> Though it be too obvious to escape observation, that different ideas are connected together; I do not find that any philosopher has attempted to enumerate or class all the principles of association; a subject, however, that seems worthy of curiosity. To me, there appear to be only three principles of connexion among ideas, namely, Resemblance, Contiguity in time or place, and Cause or Effect.[80]

The *Abstract to a Treatise of Human Nature* closes with Hume maintaining that these three relations of resemblance, contiguity and causation, are "the only ties of our thought, they are really *to us* the cement of the universe."[81] Thus Hume rules out any possibility of perceiving causation.

Sensationalism functions implicitly in yet another way in Hume's argument against causal realism, as it underwrites Hume's argument against "abstract ideas."[82] If his positive view of the nature of ideas is combined with his negative thesis that the notion of abstract general ideas is incoherent, then we have a pre-emptive argument against the tenability of any theory involving the abstract concept "causal power." Philosophers and psychologists have recently subjected Hume's sensationalism and his anti-abstractionism to damaging critique, and offered theories of perception and abstraction that provide detailed alternatives.[83] To the extent that Hume's theories rest on testable empirical claims about the psychology of perception, they appear to have been eclipsed by more recent, more empirically adequate models.

Since the argument for the constant conjunction account of causality depends in no small measure on the validity of sensationalism, it is important to ascertain what Hume's argument for it is, and whether it is cogent. What reason, after all, do we have for thinking that that the immediate contents of our minds in perception are these sensory impressions of which he makes so much? The argument for sensationalism is a consequence of the argument for perceptual anti-realism, or "phenomenalism" as it is often called. Hume's argument is the classic, but fallacious, argument from perceptual relativity. First, I present the argument and how the sensationalist doctrine arises as a result, and second, identify the logical fallacy in the argument. The relevant passages are from *Enquiry XII*:

> … nothing can ever be present to the mind but an image or perception, … the senses are only the inlets through which these images are conveyed, without being able to produce any immediate intercourse between the mind and the object. The table, which we see, seems to diminish, as we remove farther from it: But the real table, which exists

---

[79] Ibid., 283.

[80] Hume, *Enquiry*, 14.

[81] Hume, *Treatise*, 662.

[82] See Hume, *Treatise*, I.I.VII for Hume's extended critique of abstract general ideas.

[83] The tradition of direct realism most fundamentally opposed to Hume and the earlier British empiricists includes: Thomas Reid, *Thomas Reid's Inquiry and Essays*, ed. Ronald E. Beanblossom and Keith Lehrer (Indianapolis: Hackett Publishing Co., 1983); James J. Gibson, *The Senses Considered as Perceptual Systems* (Boston: Houghton Mifflin Co., 1966), and idem, *The Ecological Approach to Visual Perception* (Hillsdale, NJ: Lawrence Erlbaum, 1986); David C. Kelley, *The Evidence of the Senses: A Realist Theory of Perception* (Baton Rouge: Louisiana State University Press, 1986). See David C. Kelley, "A Theory of Abstraction," Cognition 7(3&4): 329-57, and David M. Armstrong, *Universals: An Opinionated Introduction* (Boulder: Westview Press, 1989) for two provocative recent theories of abstraction and universals.

> independent of us, suffers no alteration: It was, therefore, nothing but its image, which was present to the mind. These are the obvious dictates of reason; and no man, who reflects, ever doubted that the existences, which we consider when we say, this house and that tree, are nothing but perceptions in the mind, and fleeting copies or representations of other existences …. [84]

Hume's argument can be generalized in the following way:

(P1) No external objects are things that vary in respect x relative to the perceiver.

(P2) All objects of immediate awareness in perception are things that vary in respect x relative to the perceiver.

∴ No objects of immediate awareness in perception are external objects.

After showing, in this way, what the objects of immediate awareness in perception are *not*, he has the grace to inform us what the objects of immediate awareness *are*: impressions, construed as copies or representations of something independent of acts of awareness.

Hume's sensationalism fills two theoretical gaps: (i) it tells us what the objects of immediate awareness could be like, and (ii) explains the facts of perceptual relativity—that is, explains how (P2) could be true.

The argument above is a *prima facie* valid syllogism. However, the middle term of the argument is ambiguous: what is "respect x relative to the perceiver?" It was Thomas Reid who first pointed out that Hume's argument commits the fallacy of "four terms." Reid's refutation of Hume's argument depends on the distinction between real and apparent magnitudes. The "respect x" of the major premise refers to real magnitude, while in the minor premise, "respect x" refers to apparent magnitude. Incorporating this clarification, the "syllogism" reads:

(P1') No external objects are things that vary in real magnitude relative to the perceiver.

(P2') All objects of immediate awareness in perception are things that vary in apparent magnitude relative to the perceiver.

∴ No objects of immediate awareness in perception are external objects.

The premises of the argument are now unambiguous and true, but the conclusion obviously does not follow. The argument from perceptual relativity represents an invalid form of argument. Accordingly, Hume has failed to demonstrate that no immediate objects of awareness are external objects, and so has not provided a cogent basis for thinking that phenomenalism is a more plausible analysis of perception than direct realism. Consequently, Hume has failed to provide sufficient motivation for accepting sensationalism as an account of the basic form of our sensory awareness.

If sensationalism is false, then there is no reason to reinterpret what are *prima facie* continuous transformations and processes as discontinuous, and comprised of discrete events. If processes are given in perception as being continuous, then there is no reason why we need to postulate an inscrutable substratum of cosmic glue to bond together otherwise discrete events. As a consequence, there is no sound reason for accepting Hume's reconceptualization of causality as a relation of the necessary connection *between* events. There are no interstices between events in our perception of the world, and therefore no call for the introduction of some additional mental element in order to generate the phenomenal appearance of continuity. Whereas the Humean view takes events as basic, and the appearance of continuity as a product of imaginative synthesis, I suggest

---

[84] Hume, *Enquiry*, 104.

that events are picked out from the stream of conscious experience to mark off salient changes in the world that we perceive. Insofar as Hume's argument in §IV of the *Enquiry* depends on the ontic discreteness of events, that thread of the argument now appears somewhat unraveled.

Secondly, what remains of the significance of Hume's argument in §VII (of the *Enquiry*) that impressions of necessary connections or powers cannot be had in experience? If sensationalism is false, and there are other forms of detecting causal phenomena than Hume suggests, then *even if* the conclusion is cogently drawn, it is irrelevant. The causal realist is therefore in a position to show how it is that direct experience of causality in the world is possible.

An analysis of the content of experience does not provide any good basis for sensationalism, but neither does an analysis of its form. The empiricists, including Berkeley and Locke, may have accepted sensationalism on methodological grounds, as a sort of working assumption, itself not directly susceptible of proof. David Kelley suggests that Locke believed ideas of individual qualities "are simple in the sense of being given, because they are simple in the sense of being unanalyzable."[85] Likewise, in Hume's philosophy of mind, the complex ideas are concatenations of simple ones, and the complex ideas of objects are decomposable into the simple ideas of their original sensory impressions. Indeed, such decomposability is a necessary condition of the reality of the content of the idea. One motivation for thinking that sensationalism must be true may be that if we have complex ideas of objects, and we can, in our imagination, take those ideas apart just as we can put simple ideas together, then the simple ones must be basic. But this depends for its validity on accepting another principle, *viz.*, that "whatever the mind can analyze into elements must have been synthesized by the mind out of those elements." Kelley continues:

> This was in fact a common view at the time, something Locke may have taken for granted. The science of the mind, it was supposed, should proceed as the science of matter does—discovering atoms and then explaining the properties of wholes by showing how they are compounded out of atoms.[86]

Yet, this principle is not self-evident—and pointing out the distinction between simple and complex ideas and impressions will not suffice to establish that it is true.

In any case, at this point it should be clear that the epistemological considerations that led Hume to endorse sensationalism provide essential fuel for the constant conjunction account of causation. If there is no good reason for believing the principle that "whatever the mind can analyze into elements must have been synthesized by the mind out of those elements" then that counts against the validity of sensationalism, and against the account of causation to which it carries us.

§2. The Imaginability Criterion of Possibility

The second pillar of Hume's overall argument for the constant conjunction account of causality is what I call the "imaginability criterion of possibility." (**ICP**) It takes center stage in Hume's case for inductive skepticism—that the future may not resemble the past. Fortunately for causal realism, the argument for this thesis is unsound. The inferential mistake is subtler than the one in the argument for sensationalism, and has not received much scholarly attention.

The first relevant passage appears in *Enquiry IV:*

---

[85] Kelley, "Theory of Abstraction," 53.

[86] Ibid.

> *The contrary of every matter of fact is still possible; because it can never imply a contradiction*, and is conceived by the mind with the same facility and distinctness, as if ever so conformable to reality. … Were it demonstrably false, *it would imply a contradiction, and could never be distinctly conceived by the mind*.[87]

This passage contains the two main premises of the argument. While the text following this passage implies the conclusion, an explicit statement of it appears in a parallel section of the *Treatise:*

> We can at least conceive a change in the course of nature; which sufficiently proves, that such a change is not absolutely impossible. To form a clear idea of any thing, is an undeniable argument for its possibility, and is alone a refutation of any pretended demonstration against it.[88]

The following reconstruction of Hume's argument shows it to have what appears to be a deductively valid structure:[89]

Suppose S is some matter of fact or actual state of affairs, and S' is the contrary of S, or that S' is a change in the course of nature.

(1) If the proposition affirming S' does not imply a contradiction, then S' is possible.

(2) If a proposition affirming S' implies a contradiction, then S' is inconceivable—*i.e.*, it can not be clearly and distinctly conceived.

Given these two premises, we may reason thus:

(3) If S' is not possible, then the proposition affirming S' implies a contradiction. [from (1)]

(4) If S' is not possible, then S' can not be clearly and distinctly conceived. [from (2), (3)]

Transposing (4), Hume's conclusion follows:

∴ (5) If S' can be clearly and distinctly conceived, then S' is possible. (**ICP**)

The conclusion means that for any counterfactual state-of-affairs, S,' if S' is distinctly and clearly conceivable[90] then by that fact, S' is possible. Any course of events that we *did* experience *might have been otherwise.* In other words, S' might well have actually been the case. Furthermore, any anticipation of the *future* course of events is an imaginative construct, not a rational demonstrative inference, and is therefore just as possible as any other future course of events. At this point, all Hume needs to do is fill in some examples of instances of S' that can be imagined, and draw the conclusion that they are possible, as illustrated in this notorious passage from the *Enquiry*:

> When I see, for instance, a Billiard-ball moving in a straight line towards another; even suppose motion in the second ball should by accident be suggested to me, as the result of their contact or impulse; may I not conceive, that a hundred different events might as well follow from that cause? May not both these balls remain at absolute rest? May not the first ball return in a straight line, or leap off from the second in any line or direction? All these suppositions are consistent and conceivable. Why then should we give the preference to one, which is no more consistent or conceivable than the rest?[91]

---

[87] Hume, *Enquiry*, 15-16, emphasis added.

[88] Hume, *Treatise*, 89, emphasis added.

[89] Is not the logically most compact form in which it could be represented, but it does have the virtue of facilitating detection of the error in reasoning involved.

[90] To form a clear idea in the mind of any thing is what constitutes conceiving of the thing. The suggestion is that "conceive" and "imagine" are basically the same in meaning in this context.

[91] Hume, *Enquiry*, 18-19.

What Hume has in mind in his billiard-ball example is to show that any particular imagined sequence of events is just as conceivable, and as a prediction about the future, just as arbitrary as any other. If we imagine the sequence of events that does in fact come to pass, that is a lucky guess, not a confirmation of a *rational* expectation about the future course of events. Since any alternative to the actual outcome was conceivable, any alternative to the actual outcome *would have been possible.*

What, then, are we to make of the premises of the argument? The central problem with the argument is the ambiguous employment of the term "possibility." After analyzing the premises of the argument, we can see that the sense in which possibility is employed quietly shifts from psychological possibility to logical possibility and from logical possibility to empirical possibility.

In premise (1), the sense of possibility is clearly *logical* possibility. If a proposition affirming S' is non-contradictory, then the proposition is a logically contingent. S' may or may not be the case. Now, premise (2) says that we cannot conceive of states of affairs that imply contradictions, which seems innocent enough. After all, try as we may, we cannot generate mental images of square circles, and so on. As Hume says, "whatever … can be distinctly conceived, implies no contradiction… "[92] This is the converse of (2):

(2') If S' can be clearly and distinctly conceived, then the proposition affirming S' can never imply a contradiction.

D. B. Rasmussen assures us that this is a reasonable assertion:

> If a state of affairs can be conceived (or imagined), then it is logically possible; but if it is inconceivable (or unimaginable), then such a state of affairs is not necessarily logically impossible. The reason for this one-sided relationship is simply that what is inconceivable (or unimaginable) may merely reflect the human mind's inability to grasp something and accordingly would make the standard for logical impossibility too psychologistic, while what is conceivable (or imaginable) is a legitimate benchmark for logical possibility.[93]

If the fact that I cannot conceive something as being true is a sufficient condition for its being logically impossible, it is also the case that if I cannot conceive something as being false, *that* is a sufficient condition for its being logically necessary. I grant that whether one can or cannot conceive something as true or false is too psychologistic a criterion for logical standing.

If taking inconceivability as a sufficient condition for logical impossibility/necessity is *too* psychologistic, surely taking conceivability as a sufficient condition for logical possibility/contingency is psychologistic as well. But, to suggest that the latter is *not* too psychologistic begs the question of how, in general, to demarcate whether some criterion is too psychologistic, or not! Unless we can establish an independent standard of what counts as excessively psychologistic, we cannot defend the entailment from psychological to logical possibility. We must conclude that the psychological conceivability (or inconceivability) of a state of affairs underdetermines the logical standing of the proposition affirming it. Simply pointing out the availability of instances of logical inconsistencies for which we cannot form a mental picture is not sufficient to eliminate the underdetermination.

Consider another example. Suppose I have a standard, unadulterated six-sided die in my hand, and I conceive rolling it, and seeing a *seven* turn up. It is conceivable? Yes. Is it *logically*

---

[92] Ibid., 22.

[93] Douglas R. Rasmussen, "Logical Possibility: An Aristotelian Essentialist Critique," *Thomist* 47: 522.

possible? No. To say that it is inconceivable *because* it is logically impossible either begs the question of how we establish logical impossibility in the first place, or sunders imaginability and conceivability.

In short, in premise (1) "possibility" is used in a logical sense, while in premise (2), Hume affirms a psychological sense of possibility. Since the argument for **ICP** draws on both premises, the conclusion is ambiguously drawn.

The billiard-ball example Hume uses to illustrate the implications of his conclusion point out yet another layer of ambiguity in the sense of "possible" employed. Hume asks, "May not both these balls remain at absolute rest?" in the sense of "is it not possible that both these balls may remain at absolute rest?" In the psychological sense, it is possible—who could deny that it is a possible content of imagination? If, on the other hand, it is the logical sense that is intended, then the issue is not so clear.

We can ask another question on Hume's behalf, however: "is it not empirically possible that both these balls may remain at absolute rest?" (In other words, regardless of the creative capacities of my imagination, is it at all possible that I could experience such a bizarre turn of events?) Suppose I answer "*no*, conservation laws guarantee that it is empirically impossible." Hume has a ready response. My knowledge of conservation laws is itself based on my experience of nature. The "no" answer would only be justified if I added the additional premise that "the future will be conformable the past," that is, that the conservation laws which held in the past will continue to hold in the future. Nevertheless, this assumption is not demonstrably certain, because "a change in the course of nature may occur" does not imply a contradiction. Again, Hume falls back on the non-contradictory character of denials of matters of fact, of which *all* laws of nature are exemplifications, regardless of how well confirmed they may seem to us.

If Hume's argument were clear and cogent, and he did establish that a change in the course of nature were not impossible, then the problem of induction would be afoot:

> It would be startling indeed if ice were not slippery, water froze when heated to the boiling point, and unsupported objects floated upward; but nevertheless they are not logical impossibilities. The implication of this argument for inductive inference is plain: since there are no causal necessities between objects and events, inductive inference carries no internal warrant. And that *x*'s and *y*'s have always occurred together in the past is no guarantee by any means that they always will in the future.[94]

No such problem of induction arises however, since it is not clear that the imaginability criterion of possibility is sustained by anything more substantial than Hume's fancy.

Again, the problem with Hume's **ICP** is that it conflates two senses of possibility that are in fact distinct. What is conceivable in the imagination is surely not a sufficient condition of logical possibility, even if it is a necessary condition. Yet, the sense in which a change in the course of nature is "possible" is the "real" or empirical sense of possibility, which is entailed by neither the logical nor psychological senses of possibility. The conclusion that a change in the course of nature is "possible" in the intended sense is therefore not demonstrated.

One objection to this line of reasoning might begin by suggesting that while Hume has not shown that a change in the course of nature is possible in any univocal sense, we *cannot* say that it is *not* empirically possible. The counter-objection to this is simply to point out the *ad ignorantiam* fallacy involved. Hume may still be correct that it is possible (in some sense) for cows to leap over

---

[94] Edward H. Madden, "Hume and the Fiery Furnace," *Philosophy of Science* 38 (1): 65.

the moon, or that a future change in the course of nature might make such an event possible. What I have shown is that Hume has given us no evidence on the basis of which to accept his conclusions about the real possibility of the kinds of bizarre changes in the course of nature that he describes.

The cogency of Hume's argument for the constant conjunction account of causation depends on sensationalism and the imaginability criterion of possibility both being true. In this and the previous section, I have refuted the arguments Hume gave for these principles, and thus severed key argumentative links in the Humean case for the constant conjunction account of causality. Hume's premises do not support his conclusions. By showing how the arguments fail, I have not directly argued that those conclusions are false, only that compelling arguments for them have not been offered by Hume, or other Humeans. While I believe that Hume's conclusions are in fact false, I claim only to have succeeded in showing that his theory of causality is implausible—that the assumptions upon which the truth of the theory depends (*i.e.*, sensationalism, and the assimilation of empirical to psychological possibility) are very likely mistaken. To sustain the stronger claim of outright falsehood, I would need to argue that the basic idea of the constant conjunction or regularity approach to causality is wrong-headed. That would require showing how the various recent attempts to present a modified regularity account of causality fail for essentially the same reasons as Hume's constant conjunction account does. That kind of failure would prove, beyond a reasonable doubt, that the basic idea of the "Regularity Theory" is wrong, and that it should be jettisoned in its entirety. Unfortunately, the protracted analysis that such a refutation requires would preempt the central focus of the present work, which is to argue for the plausibility of a realist alternative to the Humean account.

From a broader historical perspective, the negative thesis just presented argues, in effect, that Hume's treatment of causality represents a wrong-turn: an ontologically deflationary counter-analysis based on conceptual confusion at crucial points. On the other hand, I do not urge simply ignoring Hume, for despite his errors, modern philosophy of science is immeasurably better off for having had to wrestle with the implications of his positions. Indeed, the signs of struggle are written all over the contemporary empiricist "paradigm"—in the assumptions about what the right questions are, and the appropriate methods for addressing them.

1.4 Adequacy Criteria for Theories of Causality

The primary aim of the present work will be to demonstrate the plausibility of the realist alternatives on each of the major issues upon which realism and Humeanism disagree. The secondary aim of the thesis is to present causal realism as a coherent overall framework for the analysis of causal questions rather than as a grab bag of unrelated and possibly incompatible theses. If I can successfully show that realist positions on the fundamental questions of the theory of causality are plausible, and that taken together they are coherent, then I assume that such is sufficient to satisfy the minimal conditions of theoretical adequacy for theories of causality in general.

§1. Plausibility and Causal Theories

In order for a theory of causality to be plausible, it must offer answers to the fundamental questions about causality such that, overall, the theory is at least as likely to be true as its theoretical competitors. In order to demonstrate that causal realism is more plausible than anti-realism, I need to show how realism answers each of the fundamental questions in a way that is *at least* as plausible as anti-realism on all of the questions, and *more* plausible on some.

What are the fundamental questions? I believe that we can identify the appropriate questions

to ask if we have certain desiderata in mind for our theories. First, the theories should be comprehensive in the sense that they engage issues at a high level of generality. The fundamental questions will, therefore, be general ones. Second, the theories should be historically-sensitive, so the basic questions will be ones that raise issues of perennial philosophical concern. Lastly, the questions should be theory-neutral, so as not to bias the outcome by including or excluding questions for which one theory may have better answers than another.

In order to show how *not* to identify the questions that any plausible theory of causality should ask, consider the list of questions posed recently by Ernest Sosa and Michael Tooley as "some of the fundamental issues that need to be considered in any attempt to formulate a satisfactory account of the nature of causation."[95] They are:

1. What is there between causal laws and causal relations? In particular, are causal relations between events logically supervenient upon causal laws together with the totality of non-causal states of affairs? If not, do causal relations at least presuppose the existence of corresponding, covering laws, or, on the contrary is a singularist account of causation correct?
2. Are causal states of affairs logically supervenient upon non-causal ones?
3. If not, is an *a posteriori* reduction of causal states of affairs to non-causal ones possible? Or is a realist approach to causation correct?
4. Is it possible for causal relations to be immediately given, either in perceptual experience, or introspectively?
5. Do causal concepts need to be analyzed, or can they be taken as analytically basic? If they do stand in need of analysis, should the analysis be one that reduces causal states of affairs to non-causal ones, or should it treat causal terms as theoretical, and offer a realist account of the meaning of those terms?

This inventory of questions is curious for several reasons. First, it is simply presented as a list of "some" of the important issues. There is no indication of what the "other" important issues are. Sosa and Tooley ignore questions pertaining to the ontological status of powers and natures, and the status of natural necessity—both questions that warrant discussion on the grounds of their historical significance. Should these be regarded as some of the "other" important questions, or do we infer that the authors regard these as less important, since otherwise, they would have taken care to mention them? Second, no indication is given as to the operative criteria of adequacy informing their selection of these *as* important. Does the neglect of questions about powers and necessity entail that they are less important, or that they are irrelevant? If it is the latter, then even *if* we could interpret this list as an implicit set of adequacy criteria, it is not theory-neutral. Third, there is no suggestion of any independent criteria that would justify the claim that these represent "some of the *fundamental* (my italics) issues that need to be considered." If some issues are fundamental and others are not, does that mean that the other issues are *derivative*—that they cannot be answered adequately without adverting to conclusions antecedently established?

In §1.5 below, I pose four questions that are historically important, theory-neutral, and broad in scope—questions that any adequate general theory of causality will need to address. As I show in §1.5, both Humean causal anti-realism and my version of causal realism satisfy this criterion of adequacy.

---

[95] Ernest Sosa and Michael Tooley, introduction to *Causation*, ed. Ernest Sosa and Michael Tooley (Oxford: Oxford University Press, 1993), 5.

In order to construct an alternative to Humean causal anti-realism, I review the arguments that have been made on behalf of realist themes, and extract what is plausible in them. I play the role of a philosophical scavenger, picking up the choicest ingredients, and synthesizing the collected insights into what I regard as the most plausible case that can be developed on each of the four questions posed above. I show how these various insights can be elaborated into theses that respond in a reasonable way to the main planks of the Humean approach on a point-by-point basis.

In the interests of maintaining a focus on the constructive orientation of the project, I have had to err on the side of exclusion when faced with choices about how much exegesis and close criticism was appropriate when discussing the work of authors with whom I am in large part sympathetic. For example, while I agree for the most part with what Nancy Cartwright has to say about natures and capacities, I am suspicious of the soundness of her (nonetheless intriguing) positions on the disunity of nature and the "patchwork" conception of natural nomicity. In a similar vein, I regard Rom Harré's idea of the field of potential as an entity whose nature is identical with its powers as a compromise of the integrity of the causal powers theory he develops with E. H. Madden.[96]

§2. Foundationalism and the Coherence of Causal Realism

Over the course of the remainder of the thesis, I will not explicitly argue for the coherence of causal realism, but *show* it. If all goes well, the coherence of the theory will emerge as a byproduct of the method I use to develop the realist position. The method is designed to help address the issue priority: what question should a causal theorist begin with, and in what order should the relevant subsequent questions be addressed? Simply entering the subject at some arbitrary point and proceeding randomly is unlikely to reap much in the way of a philosophical harvest.

Suppose we begin with the question whose possible answers depend on the answers to the fewest *other questions*. When sufficient progress has been made on the first question, then the knowledge acquired from that initial inquiry can be brought to bear on the question whose answer depends on the answer to the first question, and so on. In this way we can minimize the number of unproven assumptions needed to develop the position, identify relations of conceptual dependence, and begin the analysis with a specification of the more basic. The realist position can then be developed in a progressive way, indicating how conclusions demonstrated at one stage lead to further conclusions, and how those earlier conclusions block objections that might be raised to the more derivative ones.

Beyond this point, methodological principles cannot be rigorously defended without engaging much larger questions. In particular, my suspicion is that the most plausible realist account of causality starts with the notion that causality is detectable in experience, and affirms that our causal concepts are formed by abstraction from that experience. Theories of ideas that maintain (a) that they have their roots in the evidence of the senses, and (b) are formed by abstraction from such evidence, are theories that are included in the family of theories I call "empiricist foundationalism." The satisfaction of (a) is, so far as I can see, the best way to keep the "real" in *real*ism.

Schematically, my foundationalist requirement is similar to Hume's, *i.e.*, that every meaning-

[96] The endorsement of these authors' viewpoints on specific issues should not be construed as implying broader acceptance of their overall positions. Conversely, I do not expect that these authors will agree with the characterization of causal realism that I present, although I would hope that if they were to read the present account, they would be inclined to agree with my argument that the positions that I elaborate have more merit to them than their Humean alternatives.

ful idea derives from some impression or other. I disagree with Hume's sensationalism, phenomenalism, and nominalism. Nonetheless, we share the empiricist-foundationalist intuition that our ideas must connect up with the given in our experience at some point, otherwise we will not be able to distinguish knowledge from myth, fancy and sophistry. Moreover, both within the Aristotelian tradition and the British empiricist tradition, we find attempts to explain how general ideas or general terms can function abstractly—that is, to designate kinds and open classes of particulars. Since empiricist foundationalism rejects the notion of an "innate" idea, whatever ideas we have must ultimately derive from sense-perceptual contact with the word.

To see why this epistemic commitment is important for the causal realist's metaphysical project, suppose that one of the common objections to foundationalism is true: perhaps that all perception is theory-laden, or that it by nature involves judgement, inference or computation. In that case, there can be no perceptual awareness of causality as such; all identifications of causings, or instances of causality presuppose the contribution of some high-level cognition. It would follow that our understanding of causality would have to be (at least in part) given in terms of non-causal facts, or as some theoretical relation. The most robust form of realism is one in which causal concepts are reducible to primitive causal experiences, not to non-causal states of affairs.

While I cannot defend my foundationalist preferences rigorously in this thesis, I do show (in Chapter Two) how the foundationalist methodology is brought to bear on the problem of the preconceptual origins of our causal concepts, and thereby show how to dissolve one high-level objection to foundationalism in general.

A second large issue attends consideration of one of Sosa and Tooley's questions; that is, "which are more basic: causal laws, or causal relations?" I say: causal relations. A different answer to this question is developed by Philip Kitcher,[97] wherein the order of fundamentality runs from the theoretical system, to causal laws, to causal relations. The causal structure of the world is the *end* result of the deductive systematization of our scientific knowledge of some domain, not the evidential starting point. As a matter of methodology, one needs to articulate a system of scientific explanation first, and fit the discussion of the nature and justification of causal knowledge into the broader explanatory context. If we can justify our causal beliefs in the way that Kitcher suggests, we need not go beyond systematizing descriptions of observed phenomena, to postulate claims about causal truths independent of our quest to construct an ordered conception of those phenomena. Kitcher contrasts his "top-down" unification model of explanation with Welsey Salmon's "bottom-up" causal-mechanical model, the latter of which is, I think, clearly the more compatible with causal realism, especially if a foundationalist method is involved in developing it. Apparently so does Kitcher:

> The causal approach is wedded to a strong version of realism in which the world is seen as having a structure independent of our efforts to systematize it. […] I have been trying to show that we can make sense of scientific explanation and our view of the causal structure of nature without indulging in the metaphysics.[98]

On the other hand, Salmon maintains that

> … explanatory knowledge [is] knowledge of the hidden mechanisms by which nature works. It goes beyond phenomenal descriptive knowledge into knowledge of things

[97] Philip Kitcher, "Explanatory Unification and the Causal Structure of the World," in *Scientific Explanation*, Minnesota Studies in the Philosophy of Science, vol. XII, ed. Philip Kitcher and Wesley C. Salmon (Minnesota: University of Minnesota Press, 1989).

[98] Ibid., 499-500.

> that are not open to immediate inspection. Explanatory knowledge opens up the black boxes of nature to reveal their inner workings. It exhibits the ways in which the things we want to explain come about.[99]

Salmon has suggested that these two approaches may be amenable to at least a partial reconciliation, if it can be shown how they illuminate different aspects of scientific explanation. Salmon's ontic approach is consistent with the Aristotelian "dual-focus" theory of explanation in which the embedding of particular phenomena within objective causal structure as well as the systematization of causal generalizations are important ingredients of causal explanation.

While a causal-mechanical view of scientific explanation might be able to accommodate unification,[100] it is not clear how a unificationist could assimilate the insights of the causal-mechanical theorists, given the metaphysical enmity that "top-downers" have towards objective causal structure. The distaste for the necessary metaphysics stems, it would seem, from worries about the epistemic dimension of realist approaches to causality. But, as Kitcher admits, the "strong realist can maintain, with some plausibility, that the epistemological presuppositions of such worries about causation are unfounded."[101] I concur with this analysis. I will let Kitcher introduce my next point:

> Concerns about the possibilities of knowing about kinds and causes derive, it is suggested from an oversimplified empiricist epistemology. Empiricists worry that we can never know about causes because our basic perceptual knowledge is limited to identifying certain kinds of attributes. We can see that the first billiard ball moved, that it hit the second, and that the second moved after the impact—but … we cannot see that the impact caused the motion of the second ball.[102]

In my critique of Hume's sensationalism, I already indicated that a different approach to perception was needed, one more along the lines that D. M. Armstrong takes—that you "can just see causation."[103] In Chapter Two, we will see to what extent this line of approach can be pursued. As I indicated in §1.3.1, the argument from perceptual relativity often used to support the rejection of perceptual realism (along with its cousin, the argument from illusion) is invalid, so arguments against the perception of causality based on sensationalist assumptions will also fail. If the case for the perception of causality fails, it will do so on *other* grounds.

1.5 The Cornerstones of Causal Realism

The requirements of plausibility discussed in §1.4.1 lead me to propose that there are four fundamental questions that a theory of causality ought to answer, while, for the realist, a certain order of progression in the development of the theory is recommended. The questions are, in order: (a) Can causality be detected? (b) Is causality a law-like relation? (c) Does causality involve causal powers? (d) Do causes necessitate their effects? Naturally, this is hardly a complete list of relevant questions, but I suggest that it is comprehensive in the sense that questions addressing more technical issues will acquire their significance in relation to one or more of these four.

---

[99] Salmon, "Four Decades," 182.

[100] See Kitcher, "Explanatory Unification," 496-97 for his skeptical response.

[101] Philip Kitcher, The Advancement of Science: Science Without Legend, Objectivity Without Illusion (Oxford: Oxford University Press, 1993), 170.

[102] Ibid., 170.

[103] Ibid., n 60. See also David M. Armstrong, *Universals and Scientific Realism*, vol. II (Cambridge: Cambridge University Press, 1978), 165.

As would be expected, the staunch anti-realist and the staunch realist take opposite positions on each of the four questions. There are, of course, many different logically possible semi-Humean or semi-realist positions that could be advanced, no doubt including some coherent ones. My concern is to show that for each Humean position, the realist alternative is more plausible.

In a very general way, the differences between the Humean and the Realist perspectives can be outlined with reference to their respective answers to these four questions. This preliminary statement of the differences is unqualified and categorically stated, by intention. It is important at the outset to polarize the discussion—to represent the alternatives as a contrast of opposites. Appropriate nuances, qualifications and hedges can always be added later as these first approximations are adjusted upon intensive examination of the issues. At the outset, I want to provide a frame of reference in which the cornerstones are easily discernible from the theoretical rubble.

(a) Can causality be detected?

Humean: No. Hume believed that there was no perceptible "third thing" besides the two perceptible events in a singular causal sequence in virtue of which the causal relation obtains—no perceptible causal power in the first event that makes the second necessary. Humeans nearly all agree that "particular causal sequences cannot directly and immediately be recognized."[104] What makes a sequence of events causally related is some kind of abstract relation that holds between them. This might be best construed in terms of counterfactual dependency, nomic sufficiency, nomic necessity, unconditionality, *etc.* What is common to these technical variants is the assumption of the implicit generality of singular causal claims.

Realist: Yes. Realists acknowledge at least one of two options here. In both cases, causality is associated with the *production* of action, or the phenomenon of something's *being made to happen*. (i) In cases where we perceive certain types of complex events, the events are perceived as causal phenomena, *e.g.*, a collision between two objects, and their interaction. The complex event is perceived as an exemplification of a single causal nexus, not as a series of discrete events requiring an exogenous linking factor. (ii) In cases where we detect efficacious acts of will, the complex event of the volition-producing-action is perceived as causal—as something being made to happen.

(b) Is causality a law-like relation?

Humean*:* Yes. There is a "semantical entailment between singular causal statements and lawlike statements" such that for every singular statement there is a "law whose truth is a semantically necessary condition for the singular statement's truth."[105] This is the doctrine of *semantic causal nomism*, that connects the meaning and truth conditions for statements referring to instances of causality with the obtaining of some causal law or other. Humean empiricism maintains that singular causal claims are "implicitly general" in meaning; causality is essentially a law-like relation.

Realist: No. Instances of causality have primacy over causal laws rather than the reverse. Neither the meaning of, nor the justification of any singular causal statement necessarily de-

[104] Beauchamp and Rosenberg, *Hume,* 82.

[105] Ibid., 83-4.

pends on the fact that some causal law or other obtains, even if it is the case that there is a law associated with any causal instance.

(C) Does causality involve causal powers?

Humean: No. The Humean's objection to the idea of causal power is epistemic: our idea of "causal power" fails the test of reducibility to observation. The Humean does not deny that causal powers exist; the claim is rather that since the idea of causal power has no empirical content, no sense can be made of the notion.

Realist: Yes. Powers, including capacities and liabilities, make causality possible. The realist view to be presented here identifies causality as the process of actualization or exercise of powers.

(d) Do causes necessitate their effects?

Humean: No. The same kinds of considerations that support the dismissal of causal powers also support the dismissal of causal necessity—there is no empirical content to the notion. Further, since "any thing may produce any thing,"[106] there is an ineliminable possibility that a cause which has always heretofore had a particular effect under certain conditions, might in the future under the same conditions, not have its usual effect.

Realist: Yes. The realist generally acknowledges two possible domains of necessity: (i) a conceptual necessity in the relationship between the nature of a particular and its powers, and (ii) a natural necessity in the operations of the particulars by which causal processes are instantiated in generative mechanisms. If natural necessity obtains in the world, then our confidence in the ability of science to predict and explain events and patterns in nature is metaphysically grounded.

In this chapter, I have outlined the historical context in which the philosophical issues connected with the nature of causality were originally understood, and have emphasized that the tradition of causal realism is primarily associated with Aristotelian conceptions of nature and human knowledge. The Humean theory of causation represents a significant break with that tradition, but that break was motivated primarily by qualms that Hume had with some of the doctrines of his predecessors—qualms which, upon examination, are based on logical errors as I show in §1.3. I have formulated what I take to be the fundamental questions about causality, and sketched realist answers to them. In the following four chapters, I attempt to develop the best arguments for a realist conception of causality.

[106] Hume, *Treatise*, 173.

CHAPTER TWO: DETECTING CAUSALITY

The issue of whether we can, as D. M. Armstrong says, "just see" instances of causation is a significant one for the causal realist's project. It has been maintained, for example, that we may perceive a key causing a lock to turn, a bird's landing on a tree branch causing the branch to bend, or the hot knife causing the stick of butter to be cut. Are these claims literally true? If we can perceive—or more generally, *detect*—just one case of causation (or of causality[1]), that would be logically sufficient to refute the Humean thesis that causality is not presented to us in sense-experience. That would help clear the way for the development of the realist alternative. Moreover, given that some instances of causality are detectable, such instances turn out (*i.e.*, on a realist account) to be *foundational* for our idea of causality—the primitive content of our causal concepts is supplied by direct experience of causings.[2] The position I call "causal foundationalism" explicitly links the detection of causality to this epistemological claim. Causal foundationalism is simply the thesis that the original or primitive contents of our causal concepts are various experiential "givens," while our advanced causal concepts consist of broader abstractions from sense-experience that retain the primitive content.

Whatever disagreements among causal foundationalists exist regarding the *form* in which causality is detectable, they agree that we have direct sense-experience of causality in one form or another. This chapter argues for the plausibility of two basic claims: (a) "the detectability thesis," that is, that the commitment to the existence of cases of causality-detection is defensible; (b) "causal foundational internalism," that is, that one class of such cases of causality-detection can be utilized to secure the foundations of our knowledge of causality.

Various objections to these theses are dealt with throughout the chapter, but before proceeding, there are two objections which demand immediate attention. One objection is the Humean argument that detecting causality in sense-experience is just not possible, while the other objection argues that even if it were possible, foundationalism is untenable, so the detectable instances of causality are philosophically useless.

2.1 The Possibility of Causal Foundationalism

§1. Humean Objections to Causal Detectability

In Elizabeth Anscombe's "Causality and Determination,"[3] an argument is made for the thesis that causality is detectable in experience; she argues that in the volitional performance of actions whose effects are perceptible, such as lifting, pulling, pushing and so forth, causality is manifest in the form of causal efficacy. Acts of manipulating physical objects are perceptually detectable

---

[1] The terms "causation" and "causality" are rarely explicitly distinguished in the literature on this subject. This makes for some awkward moments in scholarship, such as when one commentator is using "causation" and the other "causality" when they have the very same concept in mind. In fact there are two different concepts that are relevant here. The first term is the more narrow of the two, referring primary to particular instances of efficient causation, or mechanical causal production. The second term tends to be employed when a more general notion is called for.

[2] Whether or not a singularist approach to causality is an appropriate ingredient of the realist's position will also in part turn on the issue of the detectability of causality. This is discussed at length in Chapter Three.

[3] G. E. M. Anscombe, "Causality and Determination," in *Causation*, ed. Ernest Sosa and Michael Tooley (Oxford: Oxford University Press, 1993).

instances of causing. Her argument has since become a lightning rod for pro-Hume critics of the detectability thesis. While some of the criticism has focused on Anscombe's particular methodological and logical presumptions,[4] the criticism has broader significance, and impacts upon the plausibility of arguments for the detectability thesis more generally. To clear the ground for the defense of the detectability thesis, we need to examine the content of Anscombe's position, the Humean counterclaims, and then show how the counterclaims fail to expose any serious flaws in Anscombe's contentions.

At one point, Anscombe responds to the Humean claim[5] that causality cannot be detected in individual cases. Despite differences between my position and hers, I concur with the sentiment of her argument here:

> Someone who says this is just not going to count anything as "observation of causality." … It is argued that "all we find" is such-and-such, and it turns out that the arguer has excluded from his idea of "finding" the sort of thing he says we don't "find."[6]

Hume's argument against our ability to find any impression of power or necessary connection in our experience was premised on the theory that impressions are of a certain nature, and that the relata of the causal relation are events of impression-reception. For Hume, causal processes are discontinuous, and there is no sensible impression of necessary connection (*i.e.*, no third thing) that would bind the discrete events together. What binds discrete events together into continuous causal processes is, for us, a custom or habit of imagination—a non-cognitive process. As Kant might say, the continuity of a causal process is a product of the "synthetic activity" of the imagination, it is not given in the "manifold in intuition."[7]

Anscombe's negative thesis is that Hume's non-detectability arguments are, at root, an implication of his philosophy of mind and epistemology—his theory of impressions and ideas and the criteria of meaning based on them. We do not directly perceive entities acting, such as a billiard ball rolling across a flat, green velvet surface, according to Hume's sensationalism; what we perceive is "an impression of travel made by the successive positions of a round white patch in our visual fields…."[8] If Hume is correct in saying that we do not perceive physical objects and some of their attributes directly, it follows that we do not perceive them acting either. So, we do not perceive them causing anything in the single case. If, on the other hand, we do perceive physical objects causing things to happen in the single case, then we do perceive causality.

Hume asks the champions of detectability to "produce some instance, wherein the efficacy is plainly discoverable to the mind, and its operations obvious to our consciousness or sensation." As Anscombe says, Hume "wants us to 'produce an instance' in which *efficacy* is related to sensation as *red* is."[9] Is this demand reasonable? Should we accept this as a criterion for whether we can perceive efficacy? Efficacy is not a sensible quality, so it is unsurprising that it would fail to be discoverable by the mind, *if* the mind is only capable of detecting the presence of sensible qualities. It is, however, an empirical assumption—and a highly dubious one—that this is true. Presumably actions such as cutting and drinking are perceptible, even though they are not sensible qualities. If so, then

---

[4] I rely on the arguments presented in Beauchamp and Rosenberg, *Hume*, 82-83.

[5] It is defended in Hume, *Enquiry*, VII.II, and idem, *Treatise* I,3.2.

[6] Anscombe, "Causality and Determination," 92.

[7] Harré and Madden, *Causal Powers*, 54-5.

[8] Anscombe, "Causality and Determination," 92

[9] Ibid.

the question of whether causal efficacy is detectable can be answered pretty much as Anscombe does: "Is cutting, is drinking … not 'efficacy?'"[10]

How might a Humean respond to the thesis that causality is perceptible as efficacy when we intentionally manipulate objects? In their critique of Anscombe, Beauchamp and Rosenberg launch a multifarious response.

First, they accuse Anscombe of adopting a controversial interpretation, namely, that Hume does not think we perceive physical objects. On the contrary, *qua* philosopher, Hume's sensationalism and phenomenalism quite clearly commit him to this view, even if in a "common sense" frame of mind, he would insist upon its opposite.

Second, being able to perceive physical objects does not imply that we also perceive causality, maintain Beauchamp and Rosenberg. Strictly speaking, that is true. However, Anscombe's examples suggest that we perceive causality when we perceive physical objects *acting*. If we can perceive the actions of entities, that *does* imply that we also perceive causality.[11]

Their third argument is a "multiple lines of defense" rebuttal of Anscombe, defending Hume:

> For the purposes of Hume's argument that particular causal sequences cannot directly and immediately be recognized, all that is required is the admission that "efficacy," or power or agency or productive force, or any of the cognates of causation, is not related to sensation in the way "red" is, whatever way that may be. That these former terms are not so related to sensation is indeed a consequence of Hume's epistemology and theory of perception. It is, however, also a consequence of a large number of other such theories, including some specifically mentioned and rejected by Hume himself.[12]

Although the authors do not mention these other mystery theories, one theory that seems to fit the description is that of John Locke. The logical point, seemingly, is this: if Hume's epistemology is just one of two or more theories that imply the non-detectability of causality, then refuting Hume is insufficient to overturn the non-detectability thesis. Anscombe's hope that rejecting parts of Hume's epistemology would "cast a pall over his theory of causation" is thus logically defective. The "multiple lines of defense" rebuttal is stated more explicitly here:

> It may be true that Hume's epistemological views and his theory of meaning first led him to a regularity account of causation, but … his own arguments may consistently be expounded outside those contexts.[13]

That may be the case, but it is beside the point. What is at issue is not whether some kind of argument *other* than the one Hume gave can be used to defend the regularity theory and the non-detectability thesis, but whether *Hume's* arguments can, or whether Hume can be interpreted in such a way that they can. This argument against Anscombe is reminiscent of the argument that Fermat gave for his "last theorem."

They then shift to a fourth objection, oddly interjected into the explanation of the third argument:

> Anscombe's attack represents a common mistake among Hume's critics, who suppose that defects in one of his theories must vitiate the central and distinctive features of other parts of his philosophy.[14]

---

[10] Ibid., 93

[11] Of course, more needs to be said about how to argue that perceiving the actions of entities is possible. I discuss this briefly, below.

[12] Beauchamp and Rosenberg, *Hume,* 83.

[13] Ibid.

This ascribes to Anscombe, and all the misguided souls like her, a very holistic view of the Humean philosophy. Hume's conclusions, they suppose, stand on their own, more or less detached from the assumptions that help to justify them. On the contrary, I think Hume's argument for the constant conjunction theory has a definite, specifiable deductive structure, the outline of which I presented in Chapter One. There, I laid out Hume's actual argument in sufficient detail to show that the non-detectability argument and the associated Regularity theory are logical outcomes of his basic assumptions. Contrary to Beauchamp and Rosenberg, refuting Hume's sensationalism and other foundational epistemic commitments, if done in conjunction with careful attention to the logical structure of his arguments and sub-arguments, *is* effective at undercutting the "central and distinctive features" of at least one part of his philosophy: his theory of causality. Beachamp and Rosenberg's stance is typical of one strain of modern scholarship on causality in which not all of Hume's premises are taken for granted, but his conclusions appear to have taken on a life of their own, detached from the argumentative context in which they were originally presented. Given the rigour with which Hume, to his credit, set down his reasoning, this is a mistake.

§2. A Dilemma for Foundationalists

Anscombe's theory, a version of foundationalism, presupposes the idea that there is some basic or primitive experience of causality the main idea of which is *us making something happen*. We can know that we make things happen, since we directly experience doing so. Let us grant that "making something happen," "producing some occurrence," or "bringing about some effect" are (roughly) equivalent means of stating what amounts to an analysis of the concept of causality. What about "making," "producing," and "bringing about effects?" Are they themselves not causal notions—to be understood in terms of causality? If so, then no non-circular analysis of causality is available in which the notion is reducible to non-causal facts. Call this the "circularity objection."

If we interpret "analysis" formally, an analysis of causation will provide a putative set of conditions the joint satisfaction of which is necessary and sufficient for the truth of sentences of the form "*a* caused *b*." If the proffered analysis does not specify such conditions in a non-circular way, then the analysis is to be rejected, or the circularity must be justified by a secondary argument. If no secondary argument is forthcoming, then the term designating the relation is a candidate for elimination. If "*a* caused *b*" can only be explicated by employing roughly synonymous terms, *e.g.*, "*a* made *b* happen," *etc.*, then that explication fails to satisfy the non-circularity criterion, and should be rejected.

A foundationalist formal analysis could take the form offered by G. H. von Wright and Douglas Gasking.[15] According to Gasking, "*A* causes *B*" means that an event or state of type *B* could be produced by means of producing a prior event or state of type *A*. Von Wright focuses on singular causation rather than on the type level, when he maintains that "*p* causes *q*" means that "if I could do *p*, then I could bring about *q*." The basic idea of the von Wright-Gasking approach is that the notion of causality is associated with intervening into and manipulating the world in order to produce desired results. A statement of the causes of some type of phenomenon is analogous to a "recipe," or an effective strategy for producing a desired effect. Action that consists of manipulating states-of-affairs to produce other actions is the "primitive" notion of causation, other more sophisti-

---

[14] Ibid.

[15] This family of theories, sometimes called "The Manipulability Theory," might also include R. G. Collingwood, "Three Senses of the Word 'Cause'" in *Philosophical Problems of Causation*, ed. Tom L. Beauchamp (Encino, CA: Dickenson, 1974): 118-25.

cated notions being derived from it, according to this model of causation.

It is easy to see that the Gasking-von Wright manipulation model, presented as an analysis of the meaning of statements of the form "Φ causes Ψ,"[16] succumbs to the circularity objection. Sterling Lamprecht, another foundational internalist, concedes the difficulty:

> From the standpoint of the requirements of a formal definition all our statements about causality are faulty. We can state the genus of causality: it is a relation. But we cannot give the essential difference of the causal relation except in some question-begging synonym. We can say that causality is a *necessary* relation between cause and effect, or that it is the character of the process in which one thing *produces* another, or that it is *efficacious* control of one thing over another. These assertions are true; but they are not adequate as formal definitions.[17]

By contrast, Hume's theory satisfies the non-circularity requirement on analyses of causality. The Humean conditions for the truth of "*a* caused *b*" reduce to temporal priority, event resemblance and event contiguity[18]—all non-causal notions.

We do not have to reject the analysis simply because of the circularity; the circularity may imply that the nature of the phenomenon under investigation may be such that the demand for analysis is theoretically inappropriate. Lamprecht, for one, is not worried about the circularity of the analysis. His position is that for the pretheoretical conception of causality, the cognitive content of the concept is indicated ostensively. "We can point to the type of fact we mean."[19] But how does Lamprecht (or anyone else) justify saying that it is really *causality* that we are pointing to? If causality is detectable, then we have some means of knowing that this is so—some means of identifying its instances.

Here's the dilemma: if we know how to identify instances of causality, then we already possess the appropriate causal concepts, or we do not. If we do already possess some causal concepts, then the detectable instances of causality in sense-experience cannot provide a preconceptual foundation for it. Detecting some experiences *as causal* presupposes that we already have the appropriate causal concepts. Suppose, on the other hand, that we did not possess any causal concepts. We would have no way to distinguish causal from non-causal states of affairs, and therefore no way to identify any of our experiences *as* being of causal states of affairs (as opposed to non-causal ones). If we could not distinguish causal from non-causal states of affairs, then there would be nothing in experience that would give rise to a need for causal concepts. Again, there could be no preconceptual foundation for causal concepts; they could never "get off the ground." We could never explain the concept to someone who did not already have it, and if no one had it, no one could acquire it in the first place.

In short, the dilemma is this: if we already have causal concepts, a preconceptual foundation is unnecessary; if we do not, a preconceptual foundation is impossible. This argument suggests that the very idea of a preconceptual foundation for causal concepts is deeply problematic, and that the foundationalist project is in principle wrong-headed.

The assumption that leads to the dilemma is the notion that in order to identify instances of a type, we have to already possess the concept of that type, or conversely, that it is not possible to

---

[16] For details, see Beauchamp and Rosenberg, *Hume*, 204-5.

[17] Sterling Lamprecht, "Causality," in *The Metaphysics of Naturalism* (New York: Meredith, 1967), 132.

[18] There are also non-reductive analyses, but these have tended to be rare. See M. Tooley's account in Chapter 3.

[19] Lamprecht, "Causality," 132.

grasp instances of a type on the basis of any pre-conceptual cognition. But this assumption is false.

We have a richly developed ability to discriminate an immense variety of similarities and contrasts among things *perceptually—i.e.,* in sense-experience. As an example, consider someone from a culture who had no exposure to either classical or jazz music. We can easily imagine that such a person could still identify the symphonies of Brahms and Beethoven as being of one type, and the music of Glenn Miller and Harry James being of another. At a pre-conceptual level, the distinction might be perceived in the form of differences in the distinctness of rhythm and in the tonal qualities of the instruments used. It would therefore be possible for such a person to identify pieces of the former as being of one type, based on the similarities in their audible characteristics. It would also be possible for samples of the latter to be identified as being of another type, based on salient *differences* in audible characteristics from the first group, all without the benefit of knowledge of musical genre. Perhaps, upon hearing a variety of samples, a relatively stable structure of similarities and differences among the various pieces would emerge in the mind of the listener. If so, then the similar instances could be regarded abstractly—*i.e.,* conceptualized—as exemplifying a type, or *genre.*

If the apprehension of relationships in sense-experience without prior conceptualization or theory is possible, then that is sufficient to forestall the first objection. Of course, a full rebuttal of the anti-foundationalist objection will depend on filling out the details of an alternative account. Below (§2.5), I explain how, without causal concepts, we can nonetheless grasp certain general features of our experience of the world—facts which will at first be conceptualized in the manner of a "common sense" understanding of the facts about causality. We fully ought to expect that these primitive "common-sense" causal notions will be superseded by the demands set by philosophical criteria of theoretical adequacy. At the same time, that does not detract from the importance of isolating a pre-theoretical context as a basic frame of reference.[20]

2.2 Varieties of Causal Foundationalism

The first family of causal concepts for which a foundation in experience has been claimed consists of these members: causes, effects, causings, causation, causality. When I speak of the "detectability" of causality in connection with its foundational significance for our causal concepts, I will use "causality" in a generic sense, referring to the class of phenomena associated with these terms. The "detectability thesis" is simply the thesis that causality (or, causation) is detectable in sense-experience in some form or other. The terms used to express the form in which causality is experienced constitute a secondary class of causal concepts, including power, force, efficacy, agency, productivity, and necessity. The terms of the first set designate aspects of the objective content of our awareness when we aware of causal phenomena. The terms of second set designate aspects of their qualitative character—*i.e.,* of the way in which those phenomena appear to us.[21]

Causal foundationalists will all agree that in some way we have direct sense-experience of causality. They also all agree that the foundational sense-experiences are all experiences of action,

---

[20] I side with Thomas Reid on the issue of the relative priority of "vulgar" and philosophical notions: when philosophic reflection deepens or sharpens our understanding, it is to be welcomed; when it contradicts our common-sense grasp of the world, it is to be rejected as sophistry.

[21] This distinction between the form and content of causality-detection should be relatively unproblematic, as it mirrors a familiar distinction in visual perception. In the same way that it makes sense to say that objects are experienced as being coloured, it makes sense to say that actions are experienced as being volitionally caused.

specifically of actions being produced or *made to happen*. Where there is action, there is causality, and *vice versa*. Causal phenomena are all associated with action—the motion, alteration or transformation of entities. On the other hand, there is considerable disagreement on the issue of the *form* in which causality is supposedly detectable. Among the contested questions are: What specific kinds of experiences may legitimately be considered instances of direct, non-inferential awareness of causality? Given the possibility that there is more than one way in which causality can be detected, a second area of disagreement is opened up: Which ways are foundational in the required sense? Is there just one basic form in which causality is detectable?

Historically, causal foundationalists have tended to endorse one of two approaches, which I refer to as causal foundational internalism (**CFI**) and causal foundational externalism (**CFE**). Each thesis is a conjunction of two claims. The first of each pair is a "detectability thesis"—a claim about the form in which causality is detectable. Foundational internalism maintains that the basic experience of causality takes the form of internal awareness of volitional efficacy, *i.e.*, of a sense of *performing* efficacious acts by means of the direct or indirect exercise of the will. The foundational externalist, on the other hand, maintains that causality in its most elementary form is the phenomenon of the production of motion, something we can sometimes *see*, when physical objects interact. The second of each pair of claims is had in common, and simply says that the first of each pair is the real foundation of causal concepts—the genuine article. Specifically, the two theses are:

Causal foundational internalism (**CFI**)

1. Causality is directly detectable as an awareness of personal volitional efficacy, and
2. Aspects of such awareness constitute the original, foundational contents of our causal concepts.

Causal foundational externalism (**CFE**)

1. Causality, in the form of the production or communication of motion, is detectable by the visual system under certain conditions of simple mechanical interaction of bodies, and
2. Aspects of such awareness constitute the original, foundational contents of our causal concepts.

Both versions come in different strains, distinguished according to the cognitive and physiological mechanisms by which the awareness of causality is generated. For example, some foundational externalists hold that it is the visual perception of our own motor activity that constitutes the evidence of efficacy, whereas others suggest that it is the visual or tactile perception of physical objects' responses when we interact with them. These alternatives are, of course, not necessarily mutually exclusive. It is possible that they may both be true *and* are equally foundational. Nevertheless, as I argue below (§2.4.4) the most plausible variant identifies the primitive awareness of causality with the articular and muscular proprioception[22] of our own motor activity.

CFI was first advocated by Thomas Reid and Maine de Biran. It asserts that causality is detectable as an experience of volitional efficacy or "*effort voulu*," in Biran's phrase. It is conceivable that this version of the detectability thesis might be regarded as somewhat obscure, and that the thesis should be expressed in more "objective" terms, emphasizing the perceptible manifestations of volition, *i.e.*, in the form of action—either of our limbs, or of the objects whose motion we produce by our own action. This was essentially the suggestion of G. E. M. Anscombe's sketch of

[22] I follow the terminological usage of Gibson, *Senses Considered*, 37. The term "kinesthesis" is sometimes used as a synonym for "articular proprioception."

her version of the detectability thesis. Anscombe's thesis that causality is perceptible is a version of foundational internalism, because it affirms that the primitive content of our causal concepts is given in the awareness we have of volitional efficacy or agency when performing such actions as lifting, pulling, *etc*. It is not, however, a "pure" form of foundational internalism since the *evidence* that our volition is efficacious is afforded by the external senses, by touch and vision.

Sterling Lamprecht is one philosopher who endorsed an externalistic version of foundationalism. According to Lamprecht, "it is from the urgency of events about us, from the way things bang and bump and push and press and clash, that we get our first experience of causality and derive our first idea of causal necessity."[23] Foundational internalism is a non-starter, for if we were to go "from the psychical facts of volition to the physical thrusts of things," then "belief in causality would be a kind of lingering animistic interpretation of the material world."[24] Nevertheless, Lamprecht subscribes to **CFI**-1, since we can eventually extend the notion of causality to include the phenomena of our own mental life, as well as the foundational experiences of "bang and bump." As for purely externalistic theories of causality-detection which reject **CFI**-1, Albert Michotte and Curt J. Ducasse have offered perhaps the most complete of such accounts.

2.3 Foundational Externalism

§1. Michotte's Theory

According to Michotte, causality can indeed be detected: it can be visually perceived under specific conditions. Michotte devised a variety of intriguing experiments designed to test under what circumstances people experience "causal impressions." In 1937 Michotte presented a paper at a Yale University conference in which he

> … expressed the opinion that certain physical events give an immediate causal impression, and that one can 'see' an object *act* on another object, *produce* in it certain changes, and *modify* it in one way or another. […] The question that arises is this: when we observe these operations, is our perception limited to the impression of two movements spatially and temporally co-ordinated, such as the advance of the knife and the cutting of the bread? Or rather do we directly perceive the action as such—do we see the knife actually cut the bread?[25]

If the latter is an accurate description of what appears in perception, then the action of the knife and the cutting of the bread are not perceived as loose and separate events, but as a single event.

If we do perceive phenomena like this as exemplifying causality, our criteria for the individuation of events will need to allow for quite "fat" events—events that are longer in duration and more complex than allowed by sensationalism. The process of perceiving events will involve a process of physiological integration of sensory information which generates a more phenomenologically complex and structured but *unitary* form in which we perceive events. The complexity is both spatial and temporal. This kind of awareness of causality would imply a commitment to the perceptual integration of sensory information indicative of changes in two spatially contiguous objects—the knife and the bread, as well as the integration over time of the transformations of the objects involved in such a way as render them as a single continuous, "coordinated" causal proc-

[23] Lamprecht, "Causality," 136.

[24] Ibid., 137.

[25] A. Michotte, *The Perception of Causality* (London: Methuen, 1963), 15.

ess.[26]

Michotte's experiments were designed to reduce the amount of information supplied to observers, to isolate the simplest situations in which causality was perceived. This way, the essential factors necessary for the "causal impression" could be identified. For Michotte, the concept of the perceptual form (or "*gestalt*") of causality refers to the stimulus structure necessary for the reception of a causal impression. There are specific, experimentally verifiable conditions under which people perceive depicted mechanical interactions as exhibiting causality. The conditions come in two basic varieties: the "Entraining Effect," and the "Launching Effect."

The launching effect is observed when one object, A, traveling toward a second stationary object, B, and comes into contact with it, stops at the time of contact, and then B beings to withdraw from the now stationary A. Under certain conditions, A is perceived as *producing*, or being the cause of B's moving off. Under other conditions, the motions of the two objects are seen as distinct. In the entraining effect, one object A, traveling towards another, stationary object, B, and comes into contact with it. The two objects then conjoin and move off together; A is perceived as "transporting" B after their interaction.

The entraining effect is the more basic of the two; only one change is observed in the "radius of action" wherein the interaction between the two objects occurs. The launching effect represents a more classical "billiard ball" scenario, where there are two distinguishable changes as a result of the interaction. The characteristic mark of the launching effect is that the two interacting objects appear to become momentarily fused, and then separate from one another. The passive object (B) acquires the impulse communicated to it by the active object during this moment of fusion, and then separates from the active object (A). The "launching" of B by A is a separation event. In the case of launching-by-striking, the effect is only clearly present if the fusion stage is less than 50ms in duration, after which time, subjects report receiving a causal impression with quickly decreasing frequency. If the fusion interval lasts longer than 0.15s, the movement of B is not perceived as being produced by A; the causal impression vanishes. Other factors can contribute to the attenuation of the launching effect. If the angle of intersection between the two object trajectories deviates from 180°, or if the speed of the two object motions differs, the causal impression is attenuated. What is required for the perception of causality is the spatiotemporal integration of the two discernible movements of the two objects into a unity, so the launching of B by A is perceived as a single event.

When the conditions for the launching effect are suboptimal, weaker types of causal impression may be perceived. In particular, the "triggering effect" occurs when the active object A is seen as "setting off" or occasioning the independent movement of B, where B is perceived as becoming unstuck or as containing within it the capacity for locomotion which needs an external trigger to get it going.

Specific conditions are likewise necessary for the entraining effect to occur. In this effect, there is no perceived delay between the two objects becoming attached and the initiation of the motion of the (previously) passive object. The subsequent motions of A and B need to be spatially co-local and along the same velocity vector. It is possible for the entraining effect to happen when the two

---

[26] This position is developed by several of the papers in William H. Warren and R. E. Shaw, eds., *Persistence and Change: Proceedings of the First International Conference on Event Perception* (Hillsdale, NJ: Lawrence Erlbaum, 1985). Many of the papers in the collection take a theoretical orientation based on James J. Gibson, *The Senses Considered as Perceptual Systems* (Boston: Houghton Mifflin Co., 1966). Gibson's own view is presented in James J. Gibson, "The Problem of Event Perception," in Edward Reed and R. Jones, eds., *Reasons for Realism: Selected Essays of James J. Gibson* (Hillsdale, NJ: Lawrence Erlbaum, 1982), 203-16.

objects are not in contact with one another. *Ceteris paribus*, the "traction effect" is a species of the entraining effect where one object (the lead object) is perceived as having the (previously) passive one "in tow."[27] The entraining effect is diminished with divergence of trajectory, velocity differences and spatial separation between A and B.

In both the entraining effect and the launching effect, there is an active object which approaches a passive object and comes into contact with it (or at least close enough to it that the subsequent actions of the objects are seen as being no different than *if* they had come into contact). The two effects differ with respect to what happens after the contact occurs. What makes them similar is that the movement of the active object is *extended to* the passive object. In Michotte's terms, there is "ampliation of the movement [of the active object]." If the movement of the first object is ampliated, then it is seen as becoming assimilated to the enlarged event (which it produces) as a mereological component. In this way, the process of movement and ampliation represents a continuous process—a single event.

Michotte has established beyond a reasonable doubt that experimental subjects identify certain sequences of actions as being causal, and provided grounds for distinguishing, for various structures of events, which are perceived as exemplifying phenomenal causality and which are not. If a physical system exhibits the same structure of events as the structures that give rise to the perception of phenomenal causality, then we are detecting causality in the physical system. There is one glaring difficulty, however: how do we know that the subjects are *right?* We know that they think they perceive causality, but *do* they?

In many of his experiments, Michotte had his subjects report on whether they receive causal impressions from events involving *projected spots of light on a wall*. Even if subjects perceive certain interactions among projected light images as manifesting causality, they are receiving the phenomenal appearance of causality where really there is just the independent motion of two projected light patches. Only where physical objects are involved can causality truly be present. As Wesley Salmon argued,[28] projected light spots are "causal pseudo-processes," and likewise their interactions are not genuine causal interactions. Receiving a causal impression may be evidence that a causal interaction is present, but if both real causal interactions and pseudo-interactions can equally well give us a perception of causality, experiencing the causal impression cannot be taken as being direct evidence of real causality.

If we stipulate that what makes the causal perception veridical is that the perceived structure of events is that of a real physical system (which Michotte's experimental apparatus was supposed to model) then what is essential to the concept of causality is the nature of the entities interacting, and not the mode of their interaction. The strongest conclusion that Michotte's results warrant is this: *If* we are observing a physical system in which a *genuine causal structure* of events is occurring, *then* we are perceiving causality. The converse does not hold; the perception of events that appears to constitute a causal process does not imply that the perceived process *is* a causal process. The import of the proviso is this: the perceptual evidence of causality, construed as the receipt of a specific pattern of transformation of the visual field, must be backed by additional information concerning the physical processes involved in generating this impression. In light of the appropriate confirming information, we can *in retrospect* verify a presumptive judgment of causality detection. In other words, detectability thesis **CFE**-1 is true. But since we need other information besides the

---

[27] Michotte, *Causality*, 160.

[28] Salmon, Causal Structure of the World, ch. 5.

causal impression—information that cannot be reduced to causal impressions—in order to justify a causality ascription to an observed sequence of events, Michotte's causal impression is not foundational for the concept of causality. The Launching and Entraining Effects are not instances of causality-detection that can be used in support of the foundational externalist position.

Ironically, Michotte's theory of the causal impression is based on a model of awareness similar to Hume's. What counts as the stimulus for a causal impression is a specific kind of transformation of the visual field, but a subject can experience the appropriate transformation in the absence of genuine causality (such as when the stimulus array is a set of projected light spots). While it may very well be that repeated observation of certain kinds of simple mechanical interactions give rise to a certain archetypally *causal* interaction which can then be simulated. But what is it about the interactions *themselves* (not our responses to them) that mark them as causal (*vs.* non-causal)? To this question, Michotte has not given us an answer.

§2. Ducasse's Theory

Like Michotte, C. J. Ducasse believed that causality was detectable in visual perception. His approach to defending the thesis diverges from Michotte's, and is more complete. He also brought a heightened sensitivity to the epistemological implications of the detectability thesis, and was able to anticipate and answer certain objections to it.

Ducasse over the course of several decades advocated a partially non-Humean theory of causation. Ducasse gave us a "full theory" in the sense that he addressed all four of the central questions concerning causation. He also gave non-Humean answers to three of them. Ducasse thought that causality was in some cases directly perceivable, that causation is a relation that holds between individual events and that causes in such cases necessitate their effects. In this section, Ducasse's case for the perceptibility of causation will be analyzed, deferring to Chapter Three his critique of "causal nomism."

To establish the intuitive plausibility of his version of the detectability thesis, Ducasse used examples in which the cause and effect are almost simultaneous and the objects involved are in physical contact. These are cases in which it is virtually impossible that the perceiver could be wrong about whatever causal claim he makes. To use Ducasse's favorite example, while witnessing a beheading, it is not possible that some unseen factor other than the falling of the guillotine was causally sufficient for the head's being detached from the neck of the condemned. That being the case, the guillotine's *causing the head to be severed from the neck* of the condemned is observable *qua* causal relation between the severing and the detaching. In this case, the sufficient cause of the beheading seems quite obviously to be the guillotine's falling; any other event in the vicinity of the execution site would be causally irrelevant to the effect in question.

For this analysis to be sustained, we need some way to take the idea of the cause being the *only* event antecedent to an effect, and to somehow show that it is the only *causally* relevant event. If sole causal relevance implies causal sufficiency, it follows that someone who sees the only event antecedent to some effect sees the cause of that event. Causality would be detectable in visual perception. Ducasse's trick is to reinterpret a principle typically assumed to be a method of inductive inference into a definition of the causal relation.

J. F. W. Herschel and J. S. Mill, the great nineteenth century advocates of inductivism, formulated a principle of empirical reasoning that W. Minto dubbed the "Method of Single Difference."[29]

[29] C. J. Ducasse, "Critique of Hume's Conception of Causality," *The Journal of Philosophy 63*(6): 146. As Ducasse notes, Hume anticipated the Victorian's formulations of the rule with his own in the *Treatise.*

The Victorian methodologists conceived of the methods of experimental inquiry as supplying argument schemas for justifying causal hypotheses—they were essentially methods of discovering causal connections in experimentally controlled contexts. According to Ducasse, the Method of Single Difference contains within it the conceptual core of the notion of causation, and thus identifies the *meaning* of the concept. Thus causation is defined in terms of a "perfect" or "strict" experiment—*i.e.*, one in which all aspects of the experimental system are held fixed, and one factor is either permitted or induced to change. In a subsequent state of the system, the single difference that obtains between it and the prior state of the system constitutes the causal relation between the two changes in the system. The causal relation between two individual concrete events is a connection between cause and effect that we observe "whenever we perceive that a certain change is the *only* one to have taken place immediately before, in the immediate environment of another."[30]

According to Ducasse, causation is an irreducibly triadic relationship. As usual, causation involves the relation between a cause event and an effect event. It also involves the "state of affairs,"[31] or the "space-time environment" in which the effect is embedded. The state of the environment, of the effect (S) and the single change (C) in that environment constitute two independent factors, besides the effect (E), that must be specified in order to characterize the causal relation properly. Since Ducasse thinks that causes and effects are typically observed to be spatially and temporally contiguous with one another, S represents both *conditions* in which the cause operates to produce the effect, and the conditions in which the effect is produced by the cause. Ducasse, unlike Mill, draws a firm distinction between causes (specifically, the changes in a system) and conditions, (the static background circumstances in which the two changes occur).

A proviso like the one given for Michotte must be made for Ducasse: if we are observing a physical system in which there are only two changes occurring, such that we observe the one change and then the other as cause and effect, respectively, then we are perceiving causality.

The proviso is necessary because even if we do perceive only two changes, it is not necessarily the case that the first caused the second. We do not perceive token changes as types of "only change in the system." Is it not the case that we would have to *infer* that some event C caused another event E in the environment, because we would have to first rule out the possibility that some unperceived change, C*, in the environment was the genuine cause of E (and C was a coincidence)? Even if observing only two changes in S were *prima facie* evidence of a causal connection between them, would we not have to have conducted a fruitless exhaustive search of the environment for alternative causes before we would be warranted in concluding that the perceived cause of E was the real cause? If so, does that not show that all particular causal connections are known not by perception but only by reason? Ducasse concedes: "the possibility still always remains that we have not in a given case observed the whole of the change in that environment."[32] To adequately justify his claim that causation is in some cases observable, he needs to specify the condition(s) that must be satisfied for it. The condition is this:

(**O**) *If* the observed change C was the only change in S prior to and in the vicinity of observed event E, *then* C was observed *causing E.*

---

[30] C. J. Ducasse, "On the Nature and Observability of the Causal Relation," in *Causation*, ed. Ernest Sosa and Michael Tooley (Oxford: Oxford University Press, 1993), 132.

[31] Edward H. Madden and James Humber, "Nonlogical Necessity and C. J. Ducasse," in *Philosophical Problems of Causation*, ed. Tom L. Beauchamp (Encino, CA: Dickenson, 1974), 165; cf. C. J. Ducasse, "Nature and Observability of the Causal Relation," 133.

[32] Ibid.

In such cases, causation is observable. This subjunctive specification of the conditions of observability of causation should not be construed as stating a condition for causal knowledge. It is an implication of what Ducasse thinks is a *metaphysical* fact about the causings: they are the sort of thing that you can often simply *see*.

The conditions for observability in question are not quite so implausible as might be thought. Consider the following analogy. Suppose you are at a hockey game, and you see Wayne Gretsky score a shorthanded goal. It is then announced over the public address system that the goal constituted a new NHL record for shorthanded goals. Even though you perceived Gretsky score a record-breaking shorthanded goal, you did not perceive the goal *as a record-breaking* shorthanded goal, because you did not know at the time that Gretsky had previously tied the shorthanded-goal record for an individual player. Likewise, we can perceive causation, even though we may not perceive the two events *as* causally connected, because we cannot necessarily rule out unperceived causes operating.

The epistemological question of how one validates the antecedent of (O)--how we know whether the observed change *is the only change* remains. Ducasse's discussion is tentative and hedged.[33] Every inference that we have observed causality when we observe only two changes in a system that satisfy the definitions of cause and effect must be guarded by an assumption: that "the circumstances which were not observed remained during the investigation causally equivalent in respect of the effect investigated."[34] Ducasse calls this postulate the Postulate of the Non-Interference of Nature, a proposition which "we hope is true."[35] Ducasse's hedge could be lessened if we took steps to ascertain whether the background unobserved conditions of the "experiment" satisfied the postulate. We do not have to know that literally everything else was constant. We just have to know that everything that *might* be causally relevant to the occurrence of the effect (given past experience) was constant. In other words, to verify that the event sequence we take to be *prima facie* causal is genuinely so, the following conditions must obtain:

(**C1**) C was perceived to be the only change in S prior to E.

(**C2**) Other possible changes C*,C**, ... that have in the past caused events similar to E in environments similar to S and were at first unperceived had later been detected present using an appropriate search technique.

(**C3**) A thorough search technique is used in an attempt to identify alternative possible causes, C*, C**, ... of E.

(**C4**) The search failed to find C*, C**, ...[36]

This procedure is sufficient to virtually eliminate the chance that some event other than C was the true cause of E. This is possible because the criteria of verification can be satisfied in practice. Skeptical responses, which presuppose criteria of verification can *not* be satisfied in practice, are logically possible, *e.g.*: "one can *never* be sure that the real cause has not gone undetected, because in any given situation, it is possible that the true cause is *imperceptible*. Thus no singular causal claim can ever be justified."

---

[33] C. J. Ducasse, *Causation and the Types of Necessity* (New York: Dover, 1969), ch. IX.

[34] Ibid., 71.

[35] Ibid., 73.

[36] This is not intended to be rigourous, but just to flesh out Ducasse's suggestion about the conditions for justifying singular causal claims.

There are at least two things wrong with the skeptical response, however. First, it "proves too much." It casts doubt upon the existence of valid criteria of justification of all universal generalizations, not just ones pertinent to establishing particular causal claims. It acquires the problem of self-application, as all skeptical arguments do. Furthermore, the meaning of the notion "possible" is ambiguous. If I say "I just saw C cause E, but it is possible that I am mistaken," I mean that I am in a situation that is similar in relevant respects to other situations where I have been proven mistaken. I might have just seen E being caused by some unseen cause, C*. Possibility of error is a reflection of my epistemic situation: not all the factors relevant to the justification of my judgment are immediately perceivable; some inference is involved. However, in skeptical arguments, it is not this sense of possibility on which their cogency depends. In skeptical arguments, possibility assertions are not empirical judgments about the epistemic situation of the perceiver (or knower) but *fancies of the imagination*. "Possibility" is used as an epistemic wild card that trumps all contextual criteria of justification and ignores the epistemic situation of actual perceivers.

Despite the failure of the skeptical objection, the argument that *genuine causality* has occurred will, in most contexts, call for auxiliary inferential techniques to confirm a presumptive judgment of causality when two events appear to be the only ones in an "experiment." The only contexts in which such auxiliary inferences are unnecessary are those in which heads are sent rolling, or in maximally similar situations.

Ducasse's argument indeed reveals that **CFE**-1 is true, in the sense that it is *possible* for causality to be detected in perception. Nonetheless, perception of two events will in general be evidentially insufficient to establish a judgment of a singular causal connection between two events. We need additional information not strictly reducible to perception to rule out hidden real causes, and to rule out the chance of being fooled by a coincidence or correlation. So again, while causality is detectable in perception, perceiving only two changes is not foundational for the concept of causality.[37]

Michotte and Ducasse's theories represent foundational externalist approaches to causality, since they both argue that causality is detectable in experience, *i.e.*, in appropriate conditions of perception. Their arguments that causality can be detected in experience are difficult to rebut directly. On the other hand, if their arguments are interpreted as foundationalist accounts of causal concepts, both positions ultimately fail. In order to justify claims that causality has been observed both accounts need to invoke auxiliary assumptions that clearly have causal import. In Michotte's theory, we have to have a causal impression, *plus* antecedent knowledge that the causal impression is a veridical representation of a genuine causal process, which cannot simply be inferred from the causal impression alone. In Ducasse's theory, we have to know that the event perceived as a cause of a specific effect was in fact causally relevant for the effect, which obviously *presupposes* some conception of causality—a conception in terms of which "causal relevance" acquires its significance. It follows that simple two-event singular causal relations cannot play the required foundational role for our causal concepts. If we are to isolate those cases of causality-detection which are genuinely foundational, we will have to look elsewhere.

---

[37] Lamprecht's discussion is relevant at this point: what is *true* of the causal relation and what is the *essence* of the causal relation are two different issues. For Lamprecht, the essence of causality consists in agency producing action, even though Lamprecht agrees that causality can be detected in experience (when heads will roll, and suchlike).

2.4 Defending Foundational Internalism I: Detecting Causality

§1. Causality as Volitional Efficacy: Reid and Biran

The correct general approach to the detectability of causality and the foundations of our causal concepts was initiated by Thomas Reid (1710-1796) and Maine de Biran (1766-1824).[38] Despite their differences, and despite errors and omissions in the details of their respective accounts, there is a valuable common theme present whose implications beg to be drawn out fully.

Reid identified causality with the action produced by the exertion of power. A Reidian characterization of the notion of causality, (or "efficiency") refers to the process of bringing about an action by means of the activation of an appropriate power. I interpret Reid to be saying that our awareness of ourselves as causal agents is the awareness that some action, whatever it is, is experienced *as being the result of*, or being brought about by, the exertion of volition. That is, our actions are experienced as consequences of our efforts of will which activate the powers that, in turn, result in bodily actions of the sort that we can deliberately perform. *Prima facie*, it appears that Reid identifies acts of will as the cause of our deliberate physical behavior, but that is not what he has in mind. The claim is that our deliberate physical behaviour is caused by *us*, *by means of* the volitional exercise of what he called "*active powers*."

"Active power" is distinguished from "intellectual power" in Reid's thought; the distinction corresponds roughly to that between natural physical *vs.* innate mental abilities. An active power is the power by which human beings bring about deliberate behaviour through the exertion of volition. Reid distinguishes between the efficacious *agent* and the *act of exerting* a power, and regarded the former as the cause in relation to the act—the effect. Because we can actualize the various active powers that we possess by acts of will directed towards the achievement of deliberate ends, we are the causes of those actions that we deliberately cause. The idea here is that the agent (the cause) produces an effect (an action) *by means of* the agent's active power, whose activation is under the direct control of the mind:

> … every action produces some change, so every change must be caused by some exertion, or by the cessation of some exertion of power. That which produces a change by the exertion of its power we call the cause of that change, and the change produced, the effect of that change.[39]

There is a three-place causal relation implied here: *A* caused *y* to happen by means of *x*. The '*x*' term in this relation is, in the volitional context, active powers of persons.

The three-place causal relation does not seem liable to reduction to the usual two-place causal relation in which the relata are events. Dropping the reference to the *A* term, leaving simply '*x* caused *y*,' implies that the *x* is not the *x* of any particular entity all, like it were some sort of occult power. To say simply that '*A* caused *y*' would imply that *y* was made to happen by no specific means at all.

For Reid, the concept of power is a "relational" conception—a conception of properties of things defined in terms of the actions such things characteristically produce. We are not cognizant of the physical nature of the underlying mechanism by which our action is brought about, so a

---

[38] While John Locke identified our active powers with our capacity for self-movement, Locke did not connect the concepts of power and causality.

[39] Quoted in Keith Lehrer, *Thomas Reid* (New York: Routledge, 1989), 205.

particular kind of power can only be identified by the effect that it is able to produce.[40] We are, however, aware of our will, and we are aware that we can use our will to exercise our active powers.

The concept of "efficient" causality pertains to physical interactions among insentient bodies, whose behavior is to be understood by analogy with "agent" causality:

> … [T]he conception of an efficient cause enters into the mind, only from the early conviction we have that we are the efficients of our own voluntary actions …. [T]he notion of efficiency will be reduced to this, that it is a relation between the cause and the effect, similar to that which is between us and our voluntary actions.[41]

The idea is that the original conceptions of cause and effect derive from our understanding the productive nature of the relationship between causes and effects in the volitional context, and then are extended by analogical reasoning to non-volitional contexts. *The efficacy of the will is the basic causal fact*.

Reid is unimpressed with Hume's "constant conjunction" account of causation. He explains as follows: "a train of events following one another ever so regularly, could never lead us to the notion of a cause, if we had not, from our constitution, a conviction of the necessity of a cause to every event."[42] Yet immediately thereafter, Reid says that "of the manner in which a cause may exert its active power, we can have no conception." If one were expecting Reid to insist that we *could* have something like a perception of active power, then this admission looks like an ironic and disappointing capitulation to Hume's skeptical arguments, advanced with considerable vigour in *Enquiry VII*. There, Hume concludes that

> … our idea of power is not copied from any sentiment or consciousness of power within ourselves, when we give rise to animal motion, or apply our limbs to their proper use and office. That their motion follows the command of the will is a matter of common experience … but the power or energy by which this is effected … is unknown and inconceivable.[43]

On a closer look, however, Reid's "capitulation" is not disappointing,[44] because there is no reason to expect that we *should* be able to detect the specific means by which volitional actuation of the will brings about its correlative effects.[45] To see why Reid was right on this point, consider the following analogy.

In visual perception, we discriminate external objects by means of a complex neurophysi-

---

[40] The concept of "power" is a "relative conception," that is, one whose qualities are identified on the basis of some other thing to which they are connected.

[41] Reid, Inquiry and Essays, 308.

[42] Ibid., 306.

[43] Hume, *Enquiry*, 44.

[44] Nor is it really surprising. Reid was a substance dualist, so the question of how the mind relates to the body is even more vexing to him than it is for property dualists like myself. If substance dualism is true, the mind-body relation is essentially ineffable, and precludes the future possibility of scientists' being able to offer a complete explanation of volitional action that integrates both the neurophysiological and folk-psychological knowledge we have about it. It is unsurprising therefore that Reid thought we could have no conception "of the manner in which a cause may exert its active power." Even for substance monists, the scientific aspects of the mind-body relation have so far resisted our most diligent efforts to understand them.

[45] It is undisappointing in this specific regard. There are other reasons way Reid's theory of human agency is unsuccessful, the most important of which involves his attempt to bring divine agency into the picture.

ological process to which we have no direct cognitive access. Arguably, the best evidence that a given neurophysiological process realizes a specific cognitive function comes when an experimenter is able to correlate a subject's experiences (based on his responses) with specific neurological interventions. Our own introspective view of things "from the inside," so to speak, is considerably less informative. When we are observing something, we are attending to a distal stimulus and are focused extrospectively on an object in our visual field. While retaining an attenuated peripheral awareness of the object, we can turn the focus of our attention inward, and attend introspectively to the fact that *we are aware* of something. This "meta-awareness" gives us qualitatively the same experience whether we attend introspectively to the fact that we are imagining, remembering, judging, inferring, or performing any other cognitive task. To notice this is to affirm the transparency of our sensory and neural hardware in our perceptual systems. Mental activity happens without our knowing how. We have a self-conscious "meta-awareness" *that* we are perceiving, imagining, reasoning *etc.*, and we are aware that *that* is under our volitional control. Similarly for our physical activity: we know that we are efficacious, but we do not know what the neurophysiological bases are for our volitional capacities.

Reid uses the concept of "power" to refer to the "how," or the "means by which" action is produced. The referent of the concept is not directly reducible to given aspects of experience as cause and effect are; it is a concept whose referents involve the neurological, physiological, and anatomic bases of human action. The specific nature of these processes eludes the frailty of inner-directed awareness. Why would we expect, then, that the forms of our awareness of our internal processes should be sufficiently informative to tell us about the details of the processes happening within our bodies? To expect so would be similar to the expectation that an abdominal cramp be able to inform us of the presence of specific toxins in the digestive tract. I interpret Reid as using "power" as an abstract explanatory concept—one that denotes the means by which a cause is connected with its effect, whatever those means are.[46] Today, two centuries later, we are little further ahead on giving a detailed account of the neurophysiological processes that realize acts of intentional causation. That Reid does not indulge any hypotheses about the "secret powers" of the nervous system is consistent with his acceptance of Newton's maxim—avoid postulating hypotheses.[47]

While Maine de Biran had somewhat more affinity with Hume's "impressions and ideas" epistemology than Reid, he is a bit clearer than Reid with identifying our original impression of causation with the "efficacious experience of willing."[48] According to Biran, any attempt to explicate the meaning of causation without reference to internal experience and *effort voulu*[49] necessarily neglects the essence of the concept. He points to the directly experienced voluntary creation of bodily movement as the impression that gives rise to our idea of causality. The emphasis here is on the internal sensory experience of what volition *feels like* when it is calling our limbs to action, and it is this inner experience of volitional effort, quite independent of the action produced, that Biran identifies as the original fact of causality.

---

[46] The nature of powers is discussed in detail in Chapter 4.

[47] Larry Laudan, "Thomas Reid and the Newtonian Turn of British Methodological Thought," in *The Methodological Heritage of Newton*, ed. R. E. Butts and J. W. Davis (Toronto: University of Toronto Press, 1970), 116.

[48] Philip Hallie, *Maine de Biran: Reformer of Empiricism* (Cambridge: Harvard University Press, 1959), 89.

[49] Ibid., 188.

§2. Against the Detectability Thesis

The most common objection to **CFI** is that its detectability thesis **CFI**-1 is simply too vague. To the extent that Biran, Reid and others have not precisely articulated just what this mysterious "*effort voulu*" is supposed to be, the theory surely could be regarded as needing more eloquent exposition. But Hume and Michotte both make a stronger point: that the theory itself is not amenable to sufficiently clear exposition.

As I interpret it, the Reid-Biran theory is a foundationalist account of a family of causal concepts, including "cause," "effect," "causality," "causal power" and their cognates. Causes are identified with agents that possess active powers, and causality is identified (in its foundational sense) with the property of agency—the power or capacity of a person to produce his or her own action. The relevant powers make action possible, so they are by that fact, *causal* powers. The powers make it metaphysically possible (if not necessary, in some cases) for the will to be efficacious in bringing about the actions that the agent intends.

According to this view, the efficacy of the will is experienced in conditions when we counter met resistance to our intended actions. In a sense, the power is also being experienced—more precisely, it is the activity or *exercise* of the power that is being experienced. This appears to be the view that Hume was describing (or something rather much like it) in the *Enquiry.* He writes:

> It may be pretended, that the resistance which we meet with in bodies, obliging us frequently to exert our force, and call up all our power, this gives us the idea of force and power. It is this *nisus* or strong endeavour, of which we are conscious, that is the original impression from which this idea is copied.[50]

In this passage, Hume maintains that **CFI** is a pretense—that the idea of power could not really come from a feeling of countering resistance. Yet, in another passage, Hume appears to capitulate:

> It must, however, be confessed, that the animal *nisus,* which we experience, though it can afford no accurate precise idea of power, enters very much into that vulgar, inaccurate idea, which is formed of it.[51]

This, of course, is exactly the concession that the **CFI**-theorist needs to win the point. The fact that the consciousness of *nisus* constitutes the grounds of a merely vulgar (*i.e.,* common-sensical) notion of power is nothing that seriously challenges **CFI**.[52] According to **CFI**, we should not expect the primitive conception acquired from sense-experience to provide all there is in terms of cognitive content. On the contrary, we should expect that these "common-sense" causal notions could be made more sophisticated if various philosophical and scientific considerations are brought to bear on their meanings.

Another objection to **CFI** was advanced by Albert Michotte, in the context of his extended critique[53] of Maine de Biran's account of causality. The willed action that gives rise to the concept of causality is called "immanent" activity by Michotte. The immanent activity is attended with a "vague impression of productivity." By contrast, there are cases, "viz. launching and entraining, in

---

[50] Hume, *Enquiry*, 44, n. 28.

[51] Ibid.

[52] One way that the "imprecision" objection could weigh against the plausibility of **CFI** is if this original concept of power could not in principle be carried to a higher level of abstraction. Hume, of course, was opposed to the idea of abstract ideas, so in conjunction with an argument against the intelligibility of abstract ideas, the objection could count against **CFI**. But Hume's arguments against abstract ideas depend crucially upon the truth of sensationalism, which has already been disposed of.

[53] Michotte, *Perception of Causality*, 266ff.

which a causal impression arises, clear, genuine, and unmistakable, and the idea of cause can be derived from it by simple abstraction…."

> Biran's mistake … was to confuse immanent activity and causality, and to suppose that the *idea* of causality which we apply to the *fact* of immanent motor activity is derived from this fact. He was led to adopt this view because immanent activity, on account of its qualitative relationship with causality, invites us to make a causal interpretation, to invest what we see with a causal significance …. [54]

In contrast to the vivid, clear experience of the causal impression obtained by witness the launching effect, the vague experience of productivity felt in immanent activity pales by comparison. While Michotte is willing to grant that the *interpretation* of agency as causal is valid, he is unwilling to grant the sensation of agency any foundational role, essentially because of its vagueness.

While the "vagueness objection" to an internalistic detectability thesis has admittedly not been presented in much detail, I am willing to grant that Hume's and Michotte's worries about vagueness are not entirely specious. The vagueness objection does seem to warrant a careful rejoinder. In order to present it, more needs to be said about how causality is be detected in sense-experience, both in terms of the phenomenology and physiology of causal experience.

Another objection that could be raised against this version of foundational internalism is that it turns on an assumption about volition—namely, that it is in some sense a real faculty of consciousness. That we have at least indirect volitional control over *some* of our behavior is a fact that ought to be plainly evident to everyone. Despite this, sometimes we hear our colleagues voice suspicions about whether this apparent volitional control is illusory.

The suspicion derives from the following line of thought. Assume that monistic physicalism is true, and that therefore our volitional capacities are grounded in our neurophysiology, and that neurophysiological processes obey deterministic causal laws. If so, that reduces the freedom of the will to the status of an illusion. Others, not quite so dogmatic in their anti-volitionalism, pursue this line of thought a little further. After all, it is said, we do not really *know* how our will functions. Yet, we cannot rule out that volition is realized in a deterministic system, so it is *possible* that freedom of the will is illusory. We are ignorant of the physical bases for our volition, so a skeptical position is encouraged. How can we be *certain* that we *can* exercise our wills freely? Answer: the same way we are sure that we know how to drive an automobile without a working knowledge of the engineering behind transmission systems and internal combustion engines. *Knowledge of the existence of a capacity is not called into question by ignorance concerning its physical basis.*

To see why the contrary position is in error, consider an example from a different context. Am I rationally warranted in adopting a skeptical stance toward the existence of gravity because I am ignorant of its physical basis? I know that gravity is the capacity of material bodies to attract one another in virtue of their masses, but I do not know *how* gravity works. I do not know whether gravity is space-time curvature, or a particle-exchange phenomenon; I have never felt the presence of gravitons. Nonetheless, I am certain that there is such a thing as gravity, and I am certain that my belief has the highest degree of empirical justification. Knowledge of the existence of the capacity of matter to attract is not called into question by ignorance the physical basis of gravity. Likewise, the existence of the capacity to will the occurrence of my actions is not called into question by ignorance of the physiological basis of volition.

---

[54] Ibid., 273.

§3. The Phenomenology of Causality-Detection

In the previous subsection, I mentioned that more needed to be said about the phenomenology and the physiology of causality detection in order to respond to the vagueness objection against **CFI**. Before we take up the first of these two points in more detail, we need to forestall one possible Humean objection.

Reid and Biran identify the foundational experience of causality with our awareness of our own efficacy when we bring about physical bodily motion by means of the deliberate exercise of some physiological capacity. *Causality* itself (in this context) refers to an aspect of the action—namely, the fact of its being a consequence of a volitional act. But to claim that the action is experienced *as being causal* would mean that we are aware of the fact that the bodily motion in question is indeed a consequence of a volitional act. On the one hand, we have the qualitative character of a state of causal awareness: in Biran's terms, a sense of "*effort voulu.*" On the other hand, we have the stipulated cognitive content: an act of volition leading to the production of a physical bodily motion by means of the exercise or activation of a causal power. If we assume that, in general, simple qualia are only associated with simple intentional contents, there is an obvious problem for the Reid-Biran view: the appropriate characterization of the cognitive content of states of causality-awareness is far more complex than the qualitative character of such states.

It might seem more likely that what we experience is a succession of discrete, isolated states: the act of volition, then the sensation of effort, then the perception of motion. From this perspective, the perception of bodily motion is the basis for an inference to the efficacy of volition, not part of the total experience of causal efficacy. If this description is correct, then we are faced with the question Hume posed in *Enquiry* §VII: what is the connection between these states, and do we have an experience of *it*, whatever it is? ("No" is the expected answer here.)

On the other hand, it is hard to see why one might suppose that our experience is quantized in this way. If it is not, then Hume's question about whether we can detect the connection between the states does not come up—there just is no "between." This stance would most likely be expressed by someone who has interpreted his or her experience in accordance with a prior commitment to sensationalism. However, if sensationalism is false, as §1.3 suggests, then there is no reason to accept this interpretation of what the relevant inner goings-on are really like.

In order to deal with the contention that **CFI** is too obscure, or too vague, I suggest that a couple of exercises might be appropriate as a means of helping the reader isolate the experience I am attempting to characterize.

> *Exercise 1:* Extend and retract your left arm a few times, attending to the inner sensations within your arm, shoulder, and upper back. Now try it again with your eyes closed. You should be able to focus on the internal sensations by which you are aware of the speed and direction of your arm, as well as the angle and orientation of your arm relative to the rest of your torso. It should be evident that you, by virtue of initiating an act of will, produced the motion of your arm. You could corroborate the internally sensed motion of the arms and fingers visually.

The awareness of volitional efficacy is more intense in the cases where our bodily motions are constrained by the presence of some external object or force that resists (to some degree) our bodily movements. To see how, perform the second exercise (or imagine performing it).

> *Exercise 2:* Pick a book up off the table. Attend to the sensation of applying extra effort to lift it than would be necessary to move your arm, unimpeded through the same motion. Notice

that you need to apply pressure with your fingers upon both of the book's covers to prevent gravity from dragging it out of your hand and noisily back onto the table. Notice the resistance of the book to the pressure exerted by your fingers upon it. Hold the book at arm's length until you become aware of the slight strain in your wrist and shoulder, as well as the feeling of your bicep contracting as you bent your arm towards you. At this point, set down the book.

Every day we encounter similar circumstances—we might notice that a revolving door is particularly stubborn and we attend to the fact that we have to consciously, intentionally, exert pressure on the glass surface with our hands in order to make the device rotate at all. On a hot humid summer day, we notice that the normally well behaved drawer in our old wooden desk now refuses to slide out; now it requires some exertion to get it unstuck and retrieve the papers stored within it. The knife that we normally use to cut onions and other vegetables in the kitchen seems to have lost its sharpness; we are suddenly aware of the extra force needed to push the blade through the items to the cutting board underneath.

The exercises isolate the sensations that are specific to volitionally-actuated motor activity, and the form in which we are aware of causality in such circumstances. The awareness of causality, as such "guided perceptions" allow us to experience, is the experience of successfully willing ourselves to perform some specific behavior.

Is it just the physical consequences of volition—the movement of our arms and other parts of our bodies—that is the evidence of our agency? In that case, the detection of volitional efficacy would be contingent upon the visual perception of bodily motion. But as the second of the previous two exercises showed, volitional efficacy is detectable as an inner sensation of motion and strain as well—surely even the blind-since-birth know that they are causally efficacious. Perhaps it is the tactile response to an object being deliberately lifted, pushed, pulled, *etc.* that is the evidence of our agency. But surely we may be aware of deliberately reaching for or throwing something without receiving any tactile stimulation, as Exercise 1 showed. These considerations suggest that there is a specific physiological mechanism, not essentially related either to touch or vision, which is responsible for generating the awareness of causality, about whose nature we are now prepared to conjecture.

§4. The Sensory Mechanism of Causality-Detection

What is the physiological basis of causality-detection? What mode of consciousness is involved in it? In this section, which is in part conjectural, I propose answers to these questions, and discuss some empirical findings from recent neurophysiological studies that tend to support my conjecture. The answers are centred on the idea that the awareness of volitional efficacy is a coordinated response of two kinds of receptors in the bodily parts whose movements we summon, called "proprioceptors" by J. J. Gibson. Proprioceptors located in the muscles and tendons appear to respond to states of contraction of muscle, *i.e.,* to differential stresses and strains. They register muscular effort rather than movement. Articular proprioceptors located throughout the joints and ligaments detect the correlative movement of the limbs.

The muscular proprioceptors, which respond to differential stresses and strains, are indicative of the presence of resisting forces. In the case of an unimpeded limb movement, for example, the proprioceptive response is specific to the force necessary to overcome the limb's own inertia. In lifting, pushing and pulling, for example, friction, gravity and spring tension, among other factors, contribute to the additional intensity of the proprioceptive response, indicative of a greater degree

of stress and strain on the muscles and tendons. The articular proprioceptors, responding to the movement of the limbs, fire in parallel with the muscular proprioceptors so that the responses of the strain/stress detection and the motion detection are sent to the brain simultaneously. The overall response of the haptic system integrates the two afferent signal streams into a unitary awareness of muscular effort-producing-action: for example, the awareness of *my lifting my arm*. This hypothesis suggests that the experience of causal efficacy is only had when the cause (the initiation of an act of will) and the effect (its attendant physical outcome) are experienced as simultaneous, or nearly simultaneous.[55]

The response of the system gives us immediate, continuous physiological feedback which tells us how our will has been causally efficacious during the course of the process of muscular exertion. Such feedback from the haptic system allows us to continually regulate and adjust the effort necessary to perform acts of manipulation demanding extreme precision (such as putting or brain surgery). It may be just the activity of such feedback structures in the central nervous system which generate the unitary sense of volitional efficacy in the experience of our own agency.

This theory appears to provide a neurophysiological model for the phenomenon that while performing motor activity, we do *not* experience a sequence of events—an act of will—a detection of force—an observation of limb movement. We are not therefore in the position of the Humean, looking for the missing "connection" *between* events. The sense of causality is part of the content of the experience of volitional efficacy, and is not an ontologically separable aspect of it.

The experience of volitional efficacy presupposes that one has awareness of one's body as a spatially continuous solid over which one has motor control. If a person lacks awareness of a part of their body, we would expect that while they may be able to touch it and visually recognize it, they will not have the sense that it is continuous with the rest of the body, and therefore lack the sense that they are able to control it (even if they in fact retain that ability). Patients who have suffered brain damage in the parietal lobe areas report just this sort of "alien limb" phenomena: visual or tactile awareness of a limb that they have no proprioceptive basis for calling *theirs*, even if they can move it or feel it. It may seem spatially detached, or seem like someone else's.

The explanation for this is indicated by recent neuropsychological studies[56] which have suggested that part of the experience of volitional efficacy depends on having a "body image" or "body schema." The body schema is a 3-D spatial memory map of the body. It assists in localizing the afferent neural information from the proprioceptors to specific points within the known spatial envelope of the body, and directs the efferent neural signals issued by motor commands to the appropriate muscles. Damage to the parietal lobes where this information is stored may prevent the brain from receiving any spatial information from the upstream neural signals. Upon action, we would get no feedback through proprioception at all, and therefore no feedback about the efficacy of any actions that we might have willed to happen. Sensorimotor coordination would be somewhat impaired, but still possible by visual and tactile guidance of, for example, limb movement.

Rosenfield's "Madame W"[57] is a bizarre case of alien limb. She *sees* her right arm, but is given no normal indication through proprioceptive (or tactile) sensation that it is hers; and so she denies

---

[55] Note the parallel with Michotte's launching effect here: in both cases there is a temporal condition attached to when causality is experienced.

[56] Israel Rosenfield, *Strange, Familiar, And Forgotten: An Anatomy Of Consciousness* (New York: Knopf Publishing Group, 1992).

[57] Ibid., 57f.

that it is hers. Then, she sees herself in a mirror, and is able to recognize the topological continuity of her body and her two arms, one of which she formerly denied was hers. Focusing purely on the visual representation of herself in the mirror, she affirms that both arms are hers. When asked to look *directly* at her arm again, after having just seen her right arm in the mirror, she is unable to claim the arm as her own at that point, since again she was getting conflicting information: proprioception told her that there was no limb of hers in that place, while her eyes told her that there was in fact *an* arm in that place.

Cases like this lend scientific support to the hypothesis that the detection of volitional efficacy critically depends on proprioception. First, as we have already seen, the required detection that our volition can produce motor activity can be obtained in the absence of any external interference with our bodies giving rise to tactile stimulation, and in the absence of any visual confirmation of that activity, both of which suggest an internal source for the senses in virtue of which we detect the efficacy in question. Second, the alien limb case suggests that if we lacked proprioception due to neurological damage, and our motion were constrained by some external body, we would only be aware of the presence of something touching us, not that it was acting as a constraint upon our motor activity. The case is similar where we acquire visual confirmation of the presence of an external constraint: without the proprioceptive evidence, we would perceive the coincidental movement of a limb, and the simultaneous movement or deformation of the proximate constraining object, but not be aware of the causal connection between the two motions. We would not be aware that *we were moving* the limb. If we could see ourselves in the mirror, like Madame W., we would have the experience either that some unknown force was moving our arm (such as might happen with direct electrical stimulation of the motor cortex), or that the arm was not *ours; i.e.*, that someone else was moving it.

We would, in this case, truly be in the situation of the Humean: thoughts of volitional intent would be constantly conjoined by motions of bodily parts such as we wished they might move, but we would not experience the motions as produced by us, and would only experience them as caused by us after we had acquired the appropriate Humean "custom."

Note that the specific mode of consciousness involved in the awareness of volitional efficacy or causal power exercise is not *introspection* but *proprioception*, of which articular and muscular kinesthesis are specific forms. I do not adopt a Cartesian foundational stance, but an empiricist one. The evidence for the efficacy of volition is not to be found in the correlation between mental acts of will and tactile or visual perception, but in the direct awareness of the volitional production of physical state changes in the body.

We experience efficacious acts of will as our *causing* the willed behavior to happen. The self-evident conscious awareness of the activity of our own volition, combined with the direct proprioceptive awareness of volitional efficacy, (enhanced in the case of constrained motion) is all the evidence we need to conclude that we are causal agents, that causality is detectable in experience, and that there is a plausible scientific explanation for how such detection is possible. In turn, **CFI** can be reformulated to reflect these claims:

Causal foundational internalism (**CFI**)

1a. Causality is directly detectable as an awareness of personal volitional efficacy by means of muscular and articular proprioception attendant upon the encounter of resistance to motion by or within the body, and

2. Aspects of such awareness constitute the original, foundational contents of our causal concepts.

The apparatus developed so far in this chapter allows us to strengthen the case for the plausibility of **CFI** insofar as it provides a basis for explaining the real force of the vagueness objections discussed in §2.4.2.

According to **CFI**-1a, muscular and articular proprioception constitute the physical basis of our (internal) consciousness of causality, while according to **CFE**-1, the visual system is the physical basis of our (external) consciousness of causality. Likewise, **CFE**-1 and **CFI**-1 require the functioning of different modes of attention to the available stimuli in order to detect causality. The vagueness objection to **CFI** was perhaps expressed most forcefully by Michotte *contra* Biran. Yet, whether it is the eye or an internal receptor on a joint that responds to a stimulus does not capture a distinction that carries much philosophical import.

What might account for the intolerance of competing options is this observation: the level of *discrimination* capable of being achieved by each of the modes of awareness involved is different. Kinesthetic or proprioceptive awareness is detected as a relatively vague, distributed sensation in the muscles, joints and tendons of the body; the stimulus for proprioceptive awareness is not discriminated very much at all. The haptic system is capable of generating a more discriminated awareness of its stimuli; we can discriminate more precisely when the system begins or ceases actively responding to the tactile stimulus. We can also spatially isolate the source of the stimulus—where it contacts the skin, and discriminate roughly is mass and size. Blind persons are able to discriminate features of the stimuli in the haptic field with astonishing resolution.[58] The visual system, of course, is capable of rendering very precise discriminations of the features of its stimuli. Differences in resolution among the perceptual systems is the only essentially cognitive difference. Given that all the perceptual systems have characteristic forms of awareness, what justifies adopting, *a priori*, some threshold level of discrimination at which the form acquires theoretical significance? If no answer is forthcoming, there is no good reason for thinking that the vagueness objection has any significance. It reflects an arbitrary preference of a resolution threshold at a point conveniently above that with which proprioception delivers, and below that with which vision delivers.

In this section, I defended the possibility of foundational internalism by rebutting local objections to the internalistic version of the detectability thesis, and I attempted to characterize the causal facts that we detect as well as describe the form in which we detect them. In the next section, I attempt to demonstrate the plausibility of **CFI**, and in effect, complete the Reidean project, by showing how causal concepts can arise from the preconceptual and pretheoretical "given" in experience.

2.5 Defending Foundational Internalism II: Conceptualizing Causal Facts

Clearly, causal concepts are applicable in both volitional and non-volitional contexts. I can cause a door to slam shut; so can the wind. Thus the idea that the volitional contexts are somehow privileged, or that we are aware of the action of causal processes *only* when human will is involved is to be rejected entirely. This implies that any analysis of the meaning of a causal judgment that does not capture the concept in its full generality will be an incomplete analysis.[59]

---

[58] Blind persons can discriminate among dozens of other people on the basis of facial topology.

[59] For example, insofar as the "manipulability theory" affirms that causality can be detected in acts of intentional manipulation, it is correct, but this is only *part* of an adequate account of causality, and not obviously the most important part.

What, then, is the common element possessed by both the ostensibly mechanical causation of simple physical systems (as in **CFE**-1) and the efficacious agent-causation of human beings? What, in addition, is the common element possessed by both the perceptible instances of causation (*e.g.*, the knife cutting butter), and other non-perceptible causings (*e.g.*, electricity traveling through a bulb causing light, *E. coli* in the drinking water causing diarrhea, constricted cranial blood vessels causing headaches, photosynthesis causing the production of oxygen in plants)? If these are truly all causal phenomena, there must either be a univocal concept of causality, or else there are multiple senses of causality and the diversity of phenomena are such that they are not amenable to subsumption by a single concept. If there is a univocal concept of causality, it must be because there is some characteristic(s) of causal facts in virtue of which they *are* causal. Whatever the meaning of the concept is, "causality" isolates the abstract respect in which the proprioceptively detectable, perceptually detectable, and nondetectable instances can be regarded as instances of essentially the same kind of thing.

§1. The Extension Problem

The challenge is to show how we can extend the notion of causality beyond its core meaning, established with reference to the will, to reach an abstract general notion of causality subsuming animate and inanimate causal production and causal efficacy. This is what I call the "Extension Problem." If the Extension Problem cannot be resolved adequately, then we are left with multiple detectability theses (**CFE**-1 and **CFI**-1) and no general conception with which to give them an integrated, unified treatment.

One proposed answer focuses on the role of causal powers. If what is essential to a cause *being* a cause is its having the power to produce an effect, then persons, other living things, and non-living things have powers. On whatever grounds we may extend the attribution of powers to objects besides ourselves, it will be by inference and *not* direct awareness of their activity in some form or another. What form does this inference take? Maine de Biran believed that he could supply the appropriate reasoning, starting from foundational internalist premises.

Here is his argument. We know that when we will our limbs to move, and they then do so, we are activating a potentially efficacious internal power. We also notice that at certain times, our will is insufficient to move the very same limbs. Assuming a state of normal health, when that occurs, the limbs are invariably attended with a tactile sensation. Since the tactile sensation is experienced as resisting one's efforts, we infer that the cause of the sensation is identical with that unknown factor that is resisting us. By means of this "primary inductive inference,"[60] Biran establishes that there exists some cause of the cessation of our willed bodily movement that is not within us. The tactile impression attending the frustrated will becomes a sign of that "indeterminate cause," such that whenever we feel what we take to be a surface we believe that there is some force behind it, but we are not clear as to the nature of that force.[61] In short, the tactile impression is inductive support for the hypothesis that there is an external *cause* of the resisting force—some power not within us.

The argument is an induction from the existence of certain sensations to some external cause of their occurrence—an inference to the best explanation. Since our haptic system does not detect the cause of a sensation directly, we must *infer* the existence of *some* cause, and use our imagination to fill in the missing details, such as that the seat of these powers is external objects of a certain

---

[60] Hallie, *Maine de Biran*, 99.

[61] Ibid.

physical structure.

Unfortunately, the argument resembles one of the Lockean arguments for the existence of the external world, and is susceptible to the same criticisms.[62] It shows the lengths to which one must go in order to establish the independent existence of the causes of our tactile sensations if a sensationalistic epistemology is assumed in advance, as Locke, Biran and the other traditional empiricists did.[63] That being said, there are yet more difficulties with Biran's theory of the acquisition of the concepts of cause and power.

As previously indicated, Biran thinks that the "primary fact" or "original sentiment" of the causal connection in experience is the efficacious experience of volitional action. By seeking the "original sentiment," Biran implies that he wants to be interpreted presenting a descriptive, ideogenetic[64] account of the concept. On the other hand, Biran is also interested in the epistemological question of how the concept's meaning comes to be broadened beyond *effort voulu* in a way that can legitimately be applied to external objects. But by being concerned to present the *grounds for inferring* the existence of force in external objects, Biran is not entitled to claim that he is involved in a purely descriptive project.

§2. Epistemic vs. Psychological Aspects of Concept-Formation

The Belgian psychologist Jean Piaget was wary of this tension in Biran's thought. "According to Maine de Biran," writes Piaget, "the order of succession in the conscious realisation coincides with that in the objective construction. To put it differently, force is first of all grasped in the self and then inferred in external objects."[65] Biran appears to be counting on the assumption that the objective construction of a concept's abstract content upon its perceptual basis[66] should parallel its ideogenesis. If he is right, then the epistemic justification of concepts by offering objective constructions of them is entirely parasitic upon presenting descriptions of how the concepts are acquired. Maine de Biran's objective construction of the concept of power indicates that we should have a primitive intuition of consciousness, much as Descartes suggested, and that the content of the concept becomes expanded by induction. The claims Biran makes for the *descriptive* status of his theory of the origin of the concept of power begs that it be tested.

Based on hundreds of interviews with children of various ages over several years, Piaget was able to falsify Biran's claims, while at the same time make some intriguing discoveries. He concludes: "If … we analyse the facts in the order of their appearance and genesis, and not in the

---

[62] Locke, *Essay,* bk. IV, ch. XI.

[63] Thomas Reid is the celebrated exception to this rule.

[64] Ideogenetic analysis traces the chronology of concepts within individuals, seeking to establish their earliest origins and empirical basis. It is not concerned with the epistemic justification or validation of concepts, but with their origins in experience and subsequent development. This is a slight modification of Piaget's concept of "genetic" analysis, renamed as "ideogenetic" in order to emphasize that it is a psychological term having nothing to do with the mechanism of transmission of inherited traits.

[65] Jean Piaget, *The Child's Conception of Physical Causality* (Totawa, NJ: Littlefield, Adams and Co., 1965), 127.

[66] "Objective construction" in this context is a method of epistemic justification of concepts. The objective construction of a concept is an idealized process of observation, differentiation, integration and abstraction by which the abstract content of a concept is derived from an appropriate set of perceptual data in accordance with certain normative principles of classification. It provides a method for establishing the objectivity of a concept by identifying the relations of conceptual dependency that forge a link between the concept and its perceptual basis.

fictitious order of adult introspection, we shall find that they admit of an interpretation which is the opposite of that put forward by Maine de Biran."[67] From the perspective of developmental cognitive psychology, Biran represents "dogmatic psychologism," which is "powerless" to lay the foundations of the ideas of cause and force. The analysis of the formation of our causal concepts ought to begin "not as they come to be observed in the mind of a philosopher but as they are seen to take place in the development of the child."[68]

Piaget's work reveals that the concept of power, which is our original causal concept, undergoes a course of development parallel to that of the child's concept of reality. The child is born *tabula rasa*; all of its experience is perceptual. No classifications are made—no categories are constructed. There is no distinction drawn between "internal" and "external," between the "I" and the external world. At an early stage in our lives, we confuse thoughts with the sounds we speak—thoughts travels via words, just as some children believe that words are whispered by trees, as their breath—the wind—travels among them. This view is one in which words are subsisting things and not essentially cognitive tools, *i.e.*, signs for other things.

The concepts of "reality" and "consciousness" are grasped later, after the child is able to differentiate the process of thought from the things thought *about*. The emergence of a separate concept of consciousness occurs when the child reduces thought to a little voice in the head and then dereifies it altogether. It becomes something *done*, not something that *is*. Similarly, the concept of reality presupposes the gradual bifurcation of the primordial melange of existence into the objective, external universe ("reality") and the subjective, internal world of thought ("consciousness"). As the child progressively differentiates reality and consciousness, signs become separated from things signified, and thought is internalized.

Just as reality and consciousness are intermixed throughout existence, the world in the mind of the young child is populated with things that exhibit the characteristics of conscious, living things. Clouds, rivers, the moon, rocks, trees, *etc.* have desires, goals, tendencies, even morals. Generally, children at this stage explain the motion of bodies by two sets of forces: an efficacious will, and a receptivity to the commands and desires of other inanimate objects. Causal explanation is never mechanistic, it usually invokes reasons that are like the reasons the child would give for the things he might do. According to Piaget, since the child is unable to differentiate himself as a conscious thing from other insentient things, the child is unable to differentiate conscious things from insentient things in general. He ascribes sentience to everything in the world. When the child begins develop a sense of self as the locus of conscious states in a world of (mostly) non-conscious entities, he has less basis for attributing consciousness to things in the world.

> In the measure that the child is ignorant of the existence of his own thought, he attributes life and consciousness to every object that comes his way, and in the measure that he discovers his own thought, he withdraws consciousness from the things around him.[69]

As this process develops, he ascribes fewer internal powers and forces to things, until gradually this tendency is abandoned completely and movement is explained in terms of mechanical impulse transference. Animism and dynamism give way to mechanism. Power is initially ubiquitous for us, and power ascriptions are restricted to an ever-narrowing range of objects as our outlook on the

---

[67] Piaget, Child's Conception of Causality, 129.

[68] Ibid., 128.

[69] Ibid., 129.

world matures. Likewise, "force is gradually withdrawn from external objects and confined within the ego." The ideogenesis of the concept of power is the complete opposite of Maine de Biran's attempted objective construction of it.

There are two lessons we can glean from Piaget's critique of Biran's theory of the formation of causal concepts. The first lesson is that an "armchair" argument for how we might intuitively think that people should form a concept may often be contradicted by careful developmental study of how people in fact do acquire concepts. Secondly, even if such an "armchair" reconstruction *is* descriptively accurate, it will not in general tell us anything about whether the process of concept formation was an "objective" construction, that is, whether the resulting notions are epistemically justified.[70]

Despite his criticisms, Piaget is still convinced that "the fact that the idea of force owes its existence to inner experience seems to be beyond dispute."[71] Piaget and Biran thus both subscribe to foundational internalism. "The idea of force," Piaget explains, "… is the result of internal experience, but not of an experience which is felt as internal from the first."[72] Piaget can be interpreted as agreeing with the "primary fact" of Maine de Biran's theory, while bringing a higher degree of epistemological sophistication to the analysis of the origins of the idea of causality. But as for Biran's theory of the "induction" by which we attribute powers to entities besides ourselves, Piaget objects: "We do not … begin by discovering internal causality and then proceed to transfer it into objects. Causality is the result of a sort of bodily contact between the organism and the world, which is prior to consciousness of self…."[73] This is an intriguing suggestion, the meaning of which, unfortunately, is rendered quite obscure by Piaget himself.

The prevailing opinion appears to be that there is no convincing way to develop a sophisticated, general notion of causality within a foundational internalist framework. It is often supposed that once foundational internalism is accepted, at best, our understanding of causality in the physical world is analogical, as Reid himself thought. At worst, we are logically committed to some form of pan-psychism or animism. We are led to the projection of our experience of volitional efficacy onto objects and events in the physical world in order to make sense of the nature of causal production more generally. Harré and Madden survey the considerable discomfort that the position engenders:

> William James was caught in this predicament … but he shrank back from drawing the pan-psychical consequence. Not seeing how to reject the premise, worried about supposing a full analogy and unable to develop a limited one, James was unable to resolve his dilemma.[74]

---

[70] The lesson can be applied more broadly. We can certainly identify the origins of a person's concept of "god," or "satyr," and discover the ideogenesis of these ideas. But the fact that we can do this tells us nothing about whether the corresponding concepts are justified. Is it not therefore sufficient to validate a concept that we be able to retrace the causal history of it in a person's mind, or even a "typical" person's mind. Rather than simply take an idea, model or theory as given, and ask whether it is satisfied by, or true of, any aspect of reality, we ought to take reality as a given, and ask whether some idea or model can be constructed objectively on that basis.

[71] Piaget, Child's Conception of Causality, 126.

[72] Ibid., 130.

[73] Ibid., 272.

[74] Harré and Madden, *Causal Powers*, 58. According to Harré and Madden, several other well-known philosophers (including F. C. S. Schiller, G. F. Stout, G. H. von Wright and Charles Hartshorne) adopted foundational internalism, but could not see how to deflect the animism objection. Alfred North Whitehead

§3. An Empiricist Grounding of Causal Concepts

In my judgement, the animistic implications attributed to **CFI** are not justified; Whitehead and the others gave up much too easily. If sufficiently careful attention is paid to the process by which the general features of causality in nature can be identified by abstraction, the extension problem will be readily resolved.

A promising approach to the extension problem has recently been developed by John Searle. Searle notices that his theory of "intentional causation" is similar enough to Reid's that he feels it necessary to emphasize the differences he has with Reid. The most important difference is that Reid allegedly believed that "the notion of causation is one that is derived from the observation we make of ourselves when we perform intentional actions."[75] As an interpretation of Reid, this is questionable, and Searle's position is (as I read Reid) in fact closer to his Scottish predecessor than he admits.

Searle formulates the internalist's version of the detectability thesis in a way that is different from Reid's version in language, but hardly so in content:

> … it is in the *performance* of actions [that we become aware of causation], for part of the Intentional content of the experience of acting when I perform intentional actions is that this experience causes the bodily movement. Notice that I am arguing here not merely that the concept of causation enters into the description of the action, but rather that part of the actual phenomena of the action is the experience of causation.[76]

In addition, Searle's description of the phenomenology of causality-detection echoes the presentation given in §2.4.3. If the sensory mechanism of causality-detection I have proposed is correct, then the Humean insistence that there must be something connecting events in any particular case of causality dissolves into nonsense. Searle concurs, and explains why this is so in the case of intentional causation:

> … one does not observe a "necessary connection" between events, rather, one event, *e.g.*, my experience of acting, is a causal Intentional presentation of the other event, *e.g.*, the movement of my arm, and the two together make up the composite event, my raising my arm.[77]

If an action or event can be causal on its own—*i.e.*, without reference to another action or event, that signals potential inadequacies of the conception of causality as essentially involving events, plus a relation between them. From the Humean standpoint, the idea that an action or an event can be an instance of causality might seem analogous to the idea of one hand clapping. That view, however, is not based on an appraisal of the facts of our experience, but on a prior theoretical commitment.

Searle then turns to the extension problem: "how … can we ever be justified in supposing entities other than our experiences can be causes and effect?" In addressing the extension problem, Searle is admirably aware of the trap Biran fell into, and distinguishes the epistemic and psychological aspects of concept-formation:

> … what follows is not intended as an empirical hypothesis about how causal concepts are acquired. … The point is not how we come by the belief that cause is a real relation

---

was another, but he was willing to bite the bullet and follow the animistic consequences to their bizarre theoretical ends.

[75] Searle also chooses von Wright as a foil. I believe that Searle is right about the differences between him and von Wright, but that he is partly wrong about Reid. See §2.4.1, above.

[76] John Searle, *Intentionality* (Cambridge: Cambridge University Press, 1983), 125.

[77] Ibid.

> in the real world, but how we might be *justified* in holding that belief …. [T]he question is, how can we be justified in supposing that something devoid of Intentionality can stand in the same relations that our Intentional states and events stand in?[78]

He then turns his attention at trying to clarify, to his own satisfaction, the implications of Piaget's obscure but intriguing suggestion that causality is "the result of a sort of bodily contact between the organism and the world." His answer to the extension problem is, briefly, this:

> We can extend the primitive experience of causation beyond the boundaries of our bodies by discovering manipulable causal regularities in the world. What we discover when we discover such a manipulable regularity is what we experience in the primitive experience of causality, the relation of one event making another event happen.[79]

He is careful to specify that discovering manipulable regularities is *learning* about the nature of particular things. The causal relations are not themselves regularities; a causal relation obtains between two things when the one thing makes the other thing happen. "Making happen" is the root notion of causality.

Searle illustrates his solution to the extension problem with the example of a child who discovers these causal facts: (a) that he can move his arm and fingers by means of exercising his will, (b) that he can pick up and move a rock by means of moving his arm and fingers, (c) that he can strike a vase by means of moving his arms, (d) he can smash a vase by means of striking it. According to Searle, this "by-means-of" relation is part of the Intentional content of the experience of causality in the volitional context. Since the "by-means-of" relation is transitive, and the "by-means-of" relation is identified as a causal relation in the first instance, causality is "carried through" each of the steps in the sequence, so that the discovery (e) that he can smash a vase by means of exercising his will in the appropriate way is also the discovery of a causal fact.

If we ask what the child has in fact discovered when he realizes that he can reliably use a rock to smash vases, it is this: a rock hitting a vase will cause it to break. A rock hitting a vase will cause it to break, whether the motion of the rock is deliberately produced with an end to breaking the vase, or whether the rock first travels through the air before hitting the vase. And if the motion of the rock by itself is sufficient (*ceteris paribus*) to break the vase, then there is no essential connection to a deliberate cause of the motion as the cause of the vase breaking. We can "detach" the original intentional action from the "by-means-of" series, so that

> … the same causation which is part of the *content* of the experience in manipulation can be *observed* in cases where there is no manipulation. The relation that the agent observes when he sees the stone smash the vase by falling on it is—as far as causation is concerned—the same relation which he experiences when he smashes the vase with the stone.[80]

This is a credible attempt at securing a basis for thinking that there is *something* called causality which is the same kind of thing in both situations, but it is ultimately unsuccessful, since it does not attempt to identify *what* causality is. Good grounds have been made for thinking that there is a common sense of "causality" involved in the two cases Searle discusses, and that manipulation is, as Piaget also suggested, the theoretical bridge that crosses the chasm of animism. To the extent that it has been shown that there is a sense in which both intentional and inanimate causality can be regarded as sharing a common element, we can regard the extension problem as solved. But it

---

[78] Ibid., 127

[79] Ibid., 132.

[80] Ibid., 129

seems like only half a solution. Even if all instances of causality involve production or "making happen," or that causes produce their effects "by means of" something or other, this is surely too abstract a description. Despite Searle's lead, the we still do not have a clear characterization of what this common element *is*.

However the extension problem is to be resolved, we will know that it has been resolved when we know how to justify ascribing powers to inanimate objects. I propose that the most promising approach to this issue begins with the recognition that our haptic system has what Reid called "a dual province." That is, on one hand, it gives us *sensory* awareness of limb movements, on the other it gives us *perceptual* awareness of objects of touch. We reject the sensationalist premise that the existence of external objects is to be inferred from tactile sensations, and so Biran's "primary induction" must be regarded as a failed attempt to establish the external existence of powerful or force-generating objects. If perceptual realism is true, the inference itself is redundant, since we can directly perceive external objects by means of the haptic system.

Awareness of the presence of any stimulus, according to J. J. Gibson, presupposes that there is some dynamic relationship between the stimulus and perceiver. If neither the perceiver nor the object of perception is in motion, consciousness is unable to extract the information in the perceptual array "afforded" by the potential object of awareness. The haptic receptors can only respond to changes in their environment; when the changes stop, soon thereafter the activity of the receptor stops and the information transmission to the brain ceases. Our awareness of the presence of the object decays and gradually drops below the level of receptor activity necessary for the awareness to occur. Can you tell just by how your head feels whether you are wearing your hat? Can you detect the presence of your undergarments? We can be totally unaware of things that are in contact with us, which implies that being in contact with the skin is not sufficient for the detection of objects. Generally, the haptic system pushes the presence of objects into conscious awareness only when there is some change in the level of the excitation of its receptors. Above a certain level of stimulation, the persistent presence of an object may remain peripherally in our consciousness. We may be constantly aware (if only peripherally) of a shoe that does not fit, because of the continual pressure against our foot.[81] Yet we can detect the presence of the slightest breeze on the face, or of a mosquito landing on the thigh. These observations imply that we can only become aware of the presence of objects if there is some force being applied, such as the force of an impact, or of a continual pressure being exerted upon us.

Human beings are not just passive perceivers. We regularly *obtain* our stimulation, by grasping, pushing, pulling and exploratory manipulation of things. The haptic system returns an enormous amount of information about the mass, structure, shape, size, texture, location, orientation, *etc.*, of objects in our environment. For example, when self-imposed tactile stimulation occurs in the grasping of an object such as a book, we impress a force upon it, we can get certain information about its physical characteristics—it is solid, inelastic, nondeformable, *hard*. No matter how firmly we grip the book, no matter how firmly our fingers press against its covers, the book will offer an equal and opposite reaction force sufficient to prevent our fingers from altering the book's macrostructure. The more firmly we grip the book, the more reaction force is exerted against our fingers. On the other hand, a soft sponge will not react in the same way, the reaction force will be

[81] Receptors must be subject to intense constant activity before they take over the central focus of our attention. But even constant intense pain begins to wane after a while—the level of intensity of the pain must be continually shifting in order for it to seize our attention.

insufficient to overcome the force of our grip, and the sponge will deform as we squeeze it.

We might be tempted to interpret the detection of the reaction force in these cases as a detection of causality. Could it be that we perceive the book *as causing* the awareness of its hardness, the sponge *as causing* the impression of its softness?

It is true that the deformation of the skin which happens when we come into contact with an object external to it is detectable by cutaneous proprioception. The question about the detection of causality is this: do we detect the sensory impressions upon our skin *as caused by the objects*? In other words, is the causality part of the content of the impression?

If we try to attend to any distinct "sensation" of touch, we will find that it is a confusion of two things: (1) an awareness of some property of the object touched, such as its hardness, smoothness, surface topology, temperature, *etc.*—in which case we have actually *perceived* some property of the object; (2) the awareness of the object as such—in which case, again, we have perceived the object. We are not aware of the object (or any of its properties) *as causing* our awareness of it. We do not detect the sensory impressions the object makes—we detect the object and some of its properties *by means of* the haptic system's overall response to the object. We are not aware of sensations *qua* receptor responses. The overall perceptual response of the haptic system is a response to the forces exerted by the objects upon our skin, which deform the skin, and receptors detect that deformation. But that is a scientific conclusion, which could not be immediately justified by the data of experience.

Reid noted the indistinctness of haptic "sensations" two centuries ago:

> … when [a man] leans his head gently against the pillar … he will tell you that he feels nothing in his head, but feels hardness in the stone. Hath he not a sensation in this case …? Undoubtedly he hath; but it is a sensation which nature intended only as a sign of something in the stone; and, accordingly, he instantly fixes his attention upon the thing signified; and cannot, without great difficulty, attend so much to the sensation as to be persuaded that there is any such thing distinct from the hardness its signifies.[82]

To the question of whether we detect the sensory impressions upon our skin *as caused by the objects*, and whether we thereby detect causality in the same manner, we must answer: no (and disagree with Searle on this point). Since we do not detect the sensory impressions themselves, there is nothing to detect *as caused*. Conversely, we do not detect the objects of awareness *as the causes* of sensations distinct from the perception of specific attributes of the objects themselves. Nor do we detect the objects *as* the causes of the perceptions. This being said, it remains true that our awareness of objects must occur *by some means*. Given realist assumptions, our perception of objects is our detection of (part of) the cause of our awareness of them. But this claim is a scientific conclusion, established by studies of physiology, anatomy and psychology.

Since we can detect, by means of our haptic systems, several different kinds of characteristics of objects by means of the obtained reaction force of objects upon us when we manipulate them, there must be some source of the force within the objects themselves. In physical terms, the hardness of an object is a function of its molecular cohesion. So it would appear that we perceive the internal (molecular) cohesion of bodies in the perceptual form of hardness. We might say that a hard body has the "power" to return a reaction force against the incident force imposed upon it by a limb in which it may be in contact—in such a way as to make the hardness perceptible. This power of the external object can be conceived as being something an object has in virtue of its

[82] Reid, Inquiry and Essays, 38.

internal structure.

As we saw earlier, we know of the *existence* of our powers of volition based on the "metaawareness" that *some* mental activity is occurring by *some* means when we are willing—in all likelihood it is a result of specific neurophysiological structure. We can justify power ascriptions to entities in general because an intrinsic structure of some kind is necessary for the possibility of the action of a thing. Whether that structure is best described in neurophysiological or molecular-structural terms, the actions that things perform are ultimately a result of the powers that the entity possesses in virtue of its intrinsic structure. To say that the powers are *causal* powers is just to indicate that their realization results in specific modes of activity—activity that is specific to the intrinsic structure that the entity has.

The similarity between the powers that inanimate objects have and the powers that human beings have lies in the explanation of how they have the powers that they do: the powers are a result of intrinsic structure.

Having outlined how the concept of "power" might first be acquired, we can also ask what features of reality are indicated by concept of "causality." In keeping with the foundationalist methodology, we need to begin by identifying what aspects of our sense-experience of the world (including ourselves, as entities in the world) give rise to our need for causal concepts. We will need to adopt a point of view deprived of our various theories and concepts about causality, and look at the world afresh, through theoretically unprejudiced eyes. In other words, we need to re-experience the world from the point of view of a person—perhaps a young child—who has not yet acquired causal concepts.

There are, I believe, two distinct sets of facts available to such a person. These facts give rise to the need for two sets of causal concepts.

Imagine a young girl playing with a favorite rag doll. At some point in her development, she will learn that she is the originator and the source of the doll's movement. This means that she learns (A) that the doll will simply lie there endlessly if she (or someone else) does not pick it up, and (B) it is *she* who makes the doll move when she picks it up. The doll does not make her move, nor does the doll "follow" her.

How does she learn (A)? Simply: by playing with the doll. After a while she will grasp that the doll is passive. Besides the fact that her doll does not move on its own, it does not speak back to her when she talks to it, regardless of whether she might *wish* it to, or whether she *imagines* that it does. If she tosses it on the ground it will not bounce up as a ball will, no matter how many times she tests whether it might.

In addition to the doll's passivity, a number of other facts are available to the young child. She will discover that some objects make noises when shaken (*e.g.,* rattles), while others do not (*e.g.,* pillows); some objects bounce when thrown or dropped, others fall to the ground and remain stationary; other objects move on their own, while others need to be pushed in order to be moved. She will notice that dogs bark, cats meow, birds chirp, *etc.* Eventually, the child will identify a pattern: that the way in which things behave is related to the type of thing that they are. What is true for dolls will be true in general, and when the child apprehends this general fact about things, she will have identified the first basic fact about causality: that things can only do or be made to do certain things and cannot do or be made to do others (no matter how much we might wish that it were in some cases otherwise.)

"Causality," in this context, names the *relation between what a thing is and how it acts.* Our knowledge of this relation is based on a generalization from all the particular facts that we learn

about how things act. Yet, we must take care to distinguish two claims:

(1) Things have specific, characteristic ways in which they can and do act.

(2) Those specific capacities for action are dependent on some aspect of the kind of thing that they are.

The first claim is a generalization from the entirety of our experience of the world; (1) is a claim about the *observed* connections between kinds of things (classified on the basis of observable characteristics) and the actions that they in fact perform. This claim expresses a pre-theoretical generalization, while (2) makes a metaphysical claim. It asserts the existence of some *explanation* for why we observe such a pattern of behavior. The explanation asserts a metaphysical dependency of the capacities of a particular upon its nature (*i.e.*, upon some of its properties).[83] *If* the realist claims can ultimately be sustained, then we ought to interpret the observed correlation between things of some type and their actions as an exemplification of an underlying ontic dependency relation between natures and capacities. But until the required argument has been made, the thesis is just a hypothesis. Indeed, at this point, even the assumption that an explanation as such is warranted has not been philosophically motivated.

The child learns (B) in much the same way that she learns (A). When she has grasped that she can manipulate the doll's movements by means of moving her arms, hands and fingers, in a repeatable, deliberate, controlled way, she conceives of herself as the origin and source of the doll's movements. She understands at that point that the particular movements the doll makes require her action and are dependent upon her. The doll will not begin moving on its own, nor will it seem to resist her.

It would be reasonable to expect that, just as fact (A) instantiates a general causal principle, fact (B) does also. The child will notice that it is not just she who can move the doll, but her friends can too, and that it is not just the doll's motion that requires her deliberate effort, but her bicycle is like that too. If this line of thought is correct, "causality," in its original sense, names the *relation between the movements of things and the things that make them move.*

[83] The argument for the realist metaphysical thesis connecting natures and capacities will be advanced in Chapters 4 and 5, below.

CHAPTER THREE: CAUSAL SINGULARISM *VS.* CAUSAL NOMISM

One of the main planks of the Humean perspective is that causality is implicitly general. Consider Hume's two definitions of "cause:"

> We may define a cause to be an object, followed by another, and where all the objects similar to the first, are followed by objects similar to the second.
>
> We may … form another definition of cause; and call it, an object followed by another, and whose appearance always conveys the thought to that other.[1]

For Hume, a "cause" is a particular defined by its relation to a universal—either a universal recurrence of objects similar to it, or universal anticipation of that recurrence. It follows that causality is that relation of one "object" to another that only obtains in the single case in virtue of the reality of some general fact of constant conjunction. The general consensus among scholars is that the "objects" so related are best understood as "events," where an event is any dated particular occurrence. As recent commentators have succinctly put the matter:

> Hume's predecessors had supposed that the causal relation was to be analyzed in terms of a particular item's inherent power, efficacy, or agency—or perhaps in the transmission of some quantity like energy, which an inherent power made possible. They also believed that causal laws are derived and established through the repeated experience of particular sequences of phenomena independently recognized as causal in character. … Hume's own brand of Copernican revolution reverses this picture: individual cases of causation are to be analyzed in terms of constant conjunctions….[2]

In this chapter, I survey the modern realist's attempt to provide a neo-classical alternative to Hume's conception of causation, and finds that a realist account is more plausible than Hume's in connection with two issues.

The primary issue concerns the validity of the Humean doctrine of "*causal nomism*." This is the doctrine according to which singular causings are implicitly general, and which affirms the primacy of causal laws. As we shall see, the argumentative support given by Hume himself for this doctrine is rather feeble. Indeed, some of Hume's apologists liken it to a postulate rather than a conclusion: "Hume's denial of the primacy of individual causal sequences in understanding causal relations is not so much a consequence of his regularity theory as its starting point."[3] Below, an alternative to causal nomism is developed in the form of a modest version of "causal singularism," which affirms the primacy of singular causal relations. This version is contrasted with Tooley's recent and quite non-Aristotelian variant.

The secondary issue involves the question of the relata of causal relations. According to general scholarly consensus, *events* are the basic relata of causal relations.[4] An alternative conception of causality is proposed which reverses the ontological hierarchy back to one more in line with the Aristotelian tradition: causal processes (entities acting in accordance with their intrinsic natures)

---

[1] Hume, *Enquiry*, 51.

[2] Rosenberg and Beauchamp, *Hume*, 80.

[3] Ibid., 81.

[4] It has also been suggested that facts can be causes in addition to events. The common feature of facts and events is that they are *dated* particulars, and it is this feature of the account of causes that I find less than adequate. If causes are entities in action, then it is the means by which a cause produces its effect that is dated; the cause itself is not.

are ontologically basic. Treating causal processes as basic has interesting implications. First, it suggests that covering laws are less important in causal explanations than the identification of the particular mechanisms by which causal processes work. Second, it accommodates a general conception of causation as a process of power-actualization. These implications are, of course, far from philosophically uncontroversial, as they involve ontological commitments to powers (including capacities and liabilities) and natures. The ontology appropriate to a realist theory of causation is dealt with in Chapter Four.

Let us first turn to the theory of causal nomism.

3.1 Causal Nomism

The Humean theory of causation gives a reductionistic account of causal relations in terms of the regular and uniform succession of tokens of event-type pairs. Two events cannot be judged to be causally linked unless a regular succession of similar event-pair instances has been observed. A causal law is the statement of such a regular succession. So knowledge of some causal law or other is a prerequisite for the justification of ascribing causal relatedness to a pair of events—without which the pair ought not to be regarded as anything other than coincidence. This epistemological doctrine of the dependence of knowledge of causality upon knowledge of causal laws is "causal nomism."

Hume repeatedly reminds us that the *knowledge* that two events are related as cause and effect cannot be acquired "*a priori*" but requires repeated, consistent observation of effects being conjoined to their causes. If this is taken to mean that causal knowledge is *a posteriori*, no realist should find offence. But Hume's claim is a stronger one: that without the relevant causal knowledge, ascriptions of causality to particular causal sequences is unwarranted. In one sense this point should be uncontroversial to the realist as well, since even in perfectly controlled experiments in which an unrepeated single trial is sufficient to deduce a causal conclusion, the realist will concede—no, insist—that such experimental inference is licensed only by a rich background of causal knowledge. Even in non-experimental contexts, "seeing" one event *as* the cause of another is often the product of tacit analogical inference. Nonetheless, the Humean claim is supposed to be quite general, and to apply even to simple cases. We could not literally "see" the motion of the knife as causing the bagel to be cut. If the Humean account were correct, we would not be entitled to infer that *this* knife is now causing *this* bagel to be cut *unless* we are able to determine some common property of knives by which they produce cuttings. The difference between causal nomism and causal singularism consists in opposite answers to the question of whether it is possible that one can know *a* caused *b*, that *a* is *F* and *b* is *G* while *not* knowing that '*F*s cause *G*s' is a law. Nomism says: no. Singularism says: yes. Schematically, nomism says that even if I perceive *a* and *b* and know that *a* is *F* and *b* is *G*:

(1) ~ know('*F*s cause *G*s' is a law) ⇒ ~know(*a* caused *b*)

or,

(2) know(*a* caused *b*) ⇒ know('*F*s cause *G*s' is a law).

Since knowledge implies truth, it follows that

(3) '*a* caused *b*' ⇒ '*F*s cause *G*s' is a law

If we streamline the symbolism somewhat, the "epistemic" aspect of causal nomism (2) as well as its "semantic" aspect (3) be stated as components of one thesis:

Epistemic Component ($\mathbf{CN}_E$)

> Any singular causal claim of the form '*c* caused *e*' depends for its justification on the knowledge of the covering causal law '*C*s cause *E*s.'[5]

Semantic Component (**CN**$_S$)

> The truth of the law '*C*s cause *E*s' is a necessary condition for the truth of the singular causal claim '*c* caused *e*.'

CN$_S$ says that for every singular causal statement, there is a causal law whose truth is a semantically necessary condition for the singular statement's truth.[6] In other words, the truth of singular causal statements depends upon the truth of a corresponding law statement. The logical relation between **CN$_E$** and **CN$_S$** is such that if **CN$_E$** is true, then **CN$_S$** will necessarily also be true. If, on the other hand, **CN$_E$** is false, the truth of **CN$_S$** is an open question.

3.2 Troubles with Causal Nomism

Perhaps the most commonly advanced consideration in favour of causal nomism is based on Hume's idea that causal relations are not apprehended directly in experience. Beginning with the ontological independence of events, according to which events are discrete, as far as we know, events can only become connected in our imagination or in thought. Insofar as causal relations are recognized at all, they are so in virtue of instantiating a law that codifies a recognized constant conjunction.

In Chapter One, I discussed the sensationalist assumptions behind Hume's case for the ontological discreteness of events and the non-detectability of causation, and found them to be unfounded. In Chapter Two, I assessed the thesis affirming the detectability of causality in experience, and found that there was a preponderance of evidence in favour of it. Taken together, these conclusions suggest that the truth of causal claims can sometimes be established by reference to immediate awareness of causality; that knowing the relevant "covering" causal laws is not a necessary condition for the justification of particular causal claims. In that case, **CN$_E$** would be false.

§1. What the Detectability Thesis Does Not *Entail*

Something like the argument just given was advanced G. E. M. Anscombe.[7] She argues, in part, that if causality can be experienced in the form of pushings, pullings, and other kinds of manipulation, then we don't need regularities to justify singular causal claims. Detecting causality provides an immediate non-inferential justification for statements about the detected facts. This has been misunderstood to imply the unanalysability of causal concepts. Before turning to the question of what the detectability thesis does entail regarding causality, we should discuss what it does not entail.

Beauchamp and Rosenberg reject Anscombe's position because

> [any] alternative to Hume's regularity approach ideally should explain why there is no directly or indirectly detectable property common and peculiar to causal sequences. Anscombe's view does not count as such an alternative, because her view is that the notion of causation is altogether unanalysable.[8]

---

[5] Here, *c* and *e* represent cause and effect event tokens, while *C* and *E* represent event types named by the existence of a property of the former token which is causally relevant to the production of the latter token.

[6] Beauchamp and Rosenberg, *Hume,* 84.

[7] Anscombe, "Causality and Determination."

[8] Beauchamp and Rosenberg, *Hume,* 83.

Sosa and Tooley, in the introduction to their anthology state that Anscombe attempts "to show that the concept of causation does not stand in need of analysis because causal relations between events can be directly perceived."[9] Since the concept of causation can be treated as basic, it is not in need of analysis in terms of other ideas.[10] The commentators agree that Anscombe is saying:

A1* Causality is directly detectable.

A2* The direct detectability of causality in the world negates the need for *any* philosophical analysis of causal concepts.

A3* A philosophical analysis of causal concepts is unnecessary and / or impossible.

Are they right?

First, we should note that Anscombe defines causality as consisting essentially "in the derivativeness of an effect from its causes."[11] One plausible reading of this, given the sort of examples she uses to illustrate causal processes, is that causality is the process of generation or production of effects by means of the action of entities. While it is a common and peculiar feature of sequences *qua* causal that they are parts of a generative process, it is true that this feature is not always observable. Some causal processes are too fast, too slow, or involve entities to small or too distant to see. So **A1*** needs to have this non-trivial qualification to be true to Anscombe's text:

A1 Causality is directly detectable in some of its instances.

If causality is only directly detectable in some of its instances, then it cannot be an essential feature of causal relations. There must be some deeper level of explanation of causality associated with the character of causal production, and the analysis of causality ought to proceed there. Given that, the imputation of **A2*** to Anscombe is mistaken. Since not *all* causal relations are detectable in immediate experience, there must be some abstract relation between stages of causal processes in terms of which the concept may to be analyzed. If this is true then clearly **A3*** is also mistaken.

The passage upon which the commentators base their assessment of Anscombe is this one:

> The truthful—though unhelpful—answer to the question: How did we come by our primary knowledge of causality? Is that in learning to speak we learned the linguistic representation and application of a host of causal concepts. Very many of them were represented by transitive and other verbs of action used in reporting what is observed. […] The word 'cause' is highly general. How does someone show that he has the concept *cause*? We may wish to say: only by having such a word in his vocabulary. If so, then the manifest possession of the concept presupposes the mastery of much else in language. I mean: the word 'cause' can be *added* to a language in which are already represented many causal concepts. A small selection: *scrape*, *push*, *wet*, *carry*, …. But if we care to imagine languages in which no special causal concepts are represented, then no description of the use of a word in such languages will be able to present it as meaning *cause*.[12]

The last sentence of this selection appears to be grist for the imputation of **A3***, since it strongly suggests the Wittgensteinian doctrine that the meaning of a term in a language is its use in the language. If that is the case, then the term "cause" has no philosophical analysis independent of its

---

[9] Sosa and Tooley, intro. to *Causation*, 13.

[10] Michael Tooley, "The Nature of Causation: A Singularist Account," *Canadian Journal of Philosophy*, Supplementary Volume 16, 286.

[11] Anscombe, "Causality and Determination," 92.

[12] Ibid., 92-3.

use as a sortal for action verbs such as those listed. Moreover, in some imaginary language in which there are no special causal verbs, the meaning of "cause" could only be grasped ostensively, since, as Anscombe maintains, the meaning of the term would be ineffable. The logical possibility of the ineffability of causality in such a language would imply that causality could not be analyzed in that language, but it hardly implies that the concept itself is incapable of analysis, as **A2*** says. In short, the textual support for A2 and A3 is present, though very weak. There is a better interpretation of Anscombe's text that is consistent with the denial of **A2*** and **A3***.

The first thing to note is that Anscombe is presenting an ideogenetic hypothesis, not a philosophical analysis of our causal concepts. As such, the thesis is supported not by formal analysis of the concept but by isolating conceptual dependencies among our concepts. What does Anscombe say about the hierarchical structure of our causal concepts? First, she identifies a list of concepts such as "pushing," "pulling" "scraping," "carrying," "eating," and so on—all implicitly causal verbs referring to the production of effects by human action. By contrasting types of motions, alterations and other changes *not* due to human action to those that are, she isolates a category of actions that have a basic common feature. The concept "causing" generalizes from these intensional actions, and the concept "cause" becomes identified with the locus of ability to make the effect happen—the locus usually being the volitional agent.

Anscombe's affirmative assessment of the detectability thesis combined with her ideogenetic hypotheses gives us the "extension problem" to deal with. But it is unreasonable to think that her view implies the in-principle unanalysability of causal concepts. Anscombe is committed to regarding *some* causal concepts as unanalysable (those whose denotation is given ostensively). That does not rule out that there are other causal concepts that are analysable *in terms of* others that are ultimately not. Indeed, a foundationalist would hold just such a claim. What her view does imply is "causal singularism"—the (epistemic) primacy of singular causal sequences.

§2. Knowing Particular Causal Relations

John Searle's rejection of causal nomism is even more explicit:

> I am walking along when suddenly a man coming the other way accidentally stumbles into me, pushing me into the gutter. Now again, barring hallucinations and the like, I know without any further observation the answer to the question "What caused you to go into the gutter?" The man bumped into me and pushed me into the gutter. In this case, one wants to say "I know all of this because I *felt* myself being pushed into the gutter, and I *saw* the man doing it to me."[13]

In this case, the causality of the process by which Searle ends up in the gutter is perceptually given—the tactile and visual evidence of the senses is sufficient in this single case to justify the claim that Searle was caused to go into the gutter by being bumped there. Similarly, Searle affirms the singularist principle that causal knowledge is possible without knowing what, if any, laws are at work:

> Suppose I am thirsty and I take a drink of water. If someone asks me why I took a drink of water, I know the answer without any further observation: I was thirsty.[14]

---

[13] Searle, *Intentionality*, 119.

[14] The example could be complicated in the direction of greater analytic precision by introducing the qualifications that the thirst did not *determine* the drinking then and there, that he had the occurrent desire to quench his thirst then, and that therefore the thirst was causally relevant but not sufficient to the occurrence of the drinking, and so on. This is a quibble, not an objection, to Searle's account.

The fact that such knowledge was available without further observation implies that any further information confirming his initial knowledge claim would be redundant. This implies specifically that knowledge pertaining to any regularities, *i.e.*, knowledge of a causal law involving thirst and water-drinking would be redundant to the justification of the knowledge claim about the causal relevance of thirst to drinking.

These considerations are sufficient to show, beyond a reasonable doubt, that $\mathbf{CN_E}$ is false. Searle's argument does not show that there is no law connecting thirst and water-drinking, but it shows that knowledge of such a law is unnecessary in order to know the truth of the assertion "My thirst caused my drinking." It still may be the case that there is a such a law.

Searle then shifts to the second issue, that is, whether one is *logically* committed to the truth of a law in virtue of his knowledge of the particular causal relation. He denies that

> "My thirst caused my drinking"

entails a statement of the form

> "there is some law *L* such that there is some description φ of my thirst and some description ψ of my drinking, and *L* asserts a universal correlation of event of type φ and events of type ψ."

In other words, he denies $\mathbf{CN_S}$, "the traditional view in its linguistic form."[15] Searle complains:

> The only argument I have ever seen is the Humean argument that since there isn't anything to causation except regularity, then for every true causal statement there must be a regularity.[16]

If saying that there's nothing to causation except regularity is the best reason for thinking that for every true causal statement there must be a regularity, then we have no *good* reasons for thinking at $\mathbf{CN_S}$ is true. Of course, it is still logically possible that $\mathbf{CN_S}$ is true (nothing that Searle says shows that $\mathbf{CN_S}$ is false); we just do not have sound reasons for believing it.

A particularly explicit discussion in which the notions of causality and nomicity are conflated appears in the section of *A System of Logic* where John Stuart Mill introduces the "Law of Universal Causation." Specifically, he identifies causality with nomic dependency:

> we recognize a law which is universal … it is co-extensive with the entire field of successive phenomena, all instances whatever of succession being examples of it. This law is the Law of Causation. The truth that every fact which has a beginning has a *cause*, is co-extensive with human experience. This generalization may appear to some minds not to amount to much, since after all it asserts only this: "it is a law, that every event depends on some *law*…"[17]

In this passage, Mill actually presents *two* statements of the Law of Causation as if they were *equivalent*.

1. Every event has a cause.
2. Every event depends on some law.

Notice that this conflation of causality and nomicity is not argued for; for Mill, the very concept of being an effect consists in its *nomic* dependence upon a cause.

Mill holds that the causal relation is characterized not just by the invariability of the antece-

---

[15] Searle, *Intentionality*, 120.

[16] Ibid.

[17] John Stuart Mill, *A System of Logic, Ratiocinative and Inductive*, 8th ed. (New York: Harper and Brothers, 1874), 236 (emphasis added).

dent to some event, but the unconditionality of the effect given the cause. In order to establish the necessity of causal relations as the basis of induction, Mill did away with the distinction between cause and condition, leaving the cause of an event as the totality of the antecedent factors invariably connected to events *of that type*. Inquiry into causes *is* inquiry into the *kinds* of antecedents that are unconditionally and invariably associated with the *kind* of effect whose causes are under investigation. C. J. Ducasse criticizes Mill's nomism with characteristic candor:

> The truth is on the contrary that it is directly and primarily an inquiry concerning *single, individual events.* If the engine of my car stops, and I ask "*Why*?" I am not asking for a statement of invariable succession or of law, even though one such may, conceivably, be inferable from the answer that it was because the magneto wire become disconnected at that moment. What I want to know is whether the latter occurrence was the single *difference* between the circumstances of the engine at the moment when it was running, and at the moment when it was not.[18]

Ducasse's critique is all the more interesting for being based on two of Mill's own ideas: the law of universal causation, and the inductive method of difference. Since all of our experience supports the thesis that no event happens without a cause, if some event happens, and the only event that happens in the otherwise stable system is a prior event, then that prior event *must* be the cause of the subsequent one. We do not have to worry about whether that kind of event happened before, nor whether it ever will again. Ducasse's argument against Mill's nomism applies at least as much to Hume himself, since it was Hume who argued for the epistemic priority of causal laws:

> Causation is *directly* concerned with single cases, not with constant conjunctions. These would follow as a matter of course, if the cause and the conditions were repeated. But constant conjunction is then a *possible corollary*, *not the definition*, of causation. To have mistaken it for the latter was Hume's epoch-making blunder, which has infected directly or indirectly, practically every discussion of causation since.[19]

Echoing Ducasse is Sterling Lamprecht:

> We can hardly suppose that single events which, taken separately, lack causal significance, can gain such significance by being grouped under a general formula. We do not reach through generalization a character which is wholly lacking in the particular facts with which we start.[20]

A similar point was made earlier by H. W. B. Joseph:

> We take uniformity in the succession of events—*i.e.*, likeness in the conditions upon which like changes succeed—to be a *sign* of causal relation, but not the same with it. For when I say that a wave striking a boat causes it to move, I imply that the relation subsists between the blow of this wave and the ensuing movement of this boat; whereas uniformity can only be exhibited in the sequence of several such movements of this or other boats … [21]

The consensus among these thinkers is that the identification of causality with either nomic regularity or nomic dependency is a mistake. Since we can know causal relations without knowing the laws with which the causal relations are supposedly connected, we cannot identify causality with nomicity in either sense. It then becomes unclear whether causality entails nomic dependency or

---

[18] Ducasse, Types of Necessity, 19.

[19] Ibid., 21.

[20] Lamprecht, "Causality," 138.

[21] H. W. B. Joseph, *Introduction to Logic*, 2d ed. (Oxford: Clarendon Press, 1914), 404.

nomic regularity. It may well be that causality *does* entail nomicity in one form or another, but it is now clear that we need an argument to the effect that such an entailment holds. In terms of the schema presented in §3.1, if $\mathbf{CN}_{\mathrm{E}}$ were true, $\mathbf{CN}_{\mathbf{S}}$ would therefore also be true. But $\mathbf{CN}_{\mathrm{E}}$ is false, so the status of $\mathbf{CN}_{\mathbf{S}}$ is, at this point, an open question.

§3. The Singularist Turn

I have argued that causal nomism's epistemic component is false, and argued in favour of the primacy of singular causes. Since epistemic causal nomism is false, the close conceptual tie between causality and nomicity is undone. Searle, Ducasse, Joseph and Lamprecht are unanimous that the conflation of causality and nomic regularity is endemic to the Humean tradition, and once these two notions are properly distinguished, the ground underneath $\mathbf{CN}_{\mathbf{S}}$ collapses. New arguments are needed if the conceptual link between causality and nomicity is to be reestablished.

Donald Davidson is one thinker who has conceded the force of the singularist critique of causal nomism.[22] Causal singularism asserts that (a) singular causal statements can be known to be true without knowing any relevant law, and (b) that singular causal statements entail no (particular) law. The latter claim is significant, for it means that I can know event c caused event e—that 'c caused e' is true—even if there is no law relating *C*s and *E*s in general. It may be the case that *C*s and *E*s exemplify no nomological relation, *even if C*-type events are regularly followed by *E*-type events. On the other hand, Davidson is clearly interested in rendering intelligible the idea of the natural world as a nomologically ordered, law-governed place, and engineers a "third way"—a reconciliation of sorts between the Humean nomological conception and the realist singularist conception.

He beings by noting an ambiguity in the Humean conception, according to which a singular causal statement '*a* caused *b*' entails that there is a law *L* such that (to paraphrase Hume) any object *x* similar to *a* is regularly succeed by some object *y* similar to *b*. The ambiguity Davidson notes is that this standard conception does not distinguish respects in virtue of which *x* and *a* are similar on the one hand, and in which *y* and *b* are similar on the other—surely some respects are causally relevant and others are not. Davidson has suggested a position that acknowledges the primacy of singular causal relations, but retains, at the same time, an attenuated form of nomicity for causality. The idea is that if *a* caused *b* then there must have been *some* description or other of *a* and *b* such that they are nomologically connected *under that description*. More formally,

$Cab \supset (\exists P)(\exists Q)(Pa \,\&\, Qa \,\&\, L)$, where

$L \equiv \Box\,[(w)(Pw \supset (\exists z)(Qz \,\&\, Swz \,\&\, z \neq w)]$[23], '$Swz$' means '$w$ is a successor of z.'

All that is required on Davidson's account is that there is some way to type-identify the token events of a singular causal relation so that the token events are backed by some law or other, even though we may never know the appropriate descriptions of the events in virtue of which they might be so backed.

We can manufacture many examples that fit Davidson's theory, but there just as many examples that disconfirm it. To take just one, suppose the causal statement in question is this: "The Chicxulub meteor impact caused the extinction of the dinosaurs." In this case, we need a rather complicated description of the cause: "the meteor impact which caused enough dust to be thrown

---

[22] Donald Davidson, "Causal Relations," in *Essays on Actions and Events* (Oxford: Clarendon Press, 1980), 160.

[23] The material conditional here is a surrogate for the elusive "causal connective."

into the atmosphere to cause sufficient sunlight to be blocked, which caused global climatic changes…. " The effect also would be have to be described in more general terms as simply, an extinction. The law would then be "whenever a meteor impact sufficient to induce global climatic changes by means of causing dust to be thrown into the atmosphere sufficient to block considerable sunlight for a period of time, then an extinction will occur." Yet it is rather unlikely that such a law is true in general. What would make such a law true? It would have to be considerably more complicated—it would have to identify and characterize all the causally relevant properties of the two events in question. Yet, for any description of the cause that is in any way general, it is still possible that the cause-type could be instantiated and yet the effect not occur, because of the extreme sensitivity of the Earth's climate pattern as a whole to small variations in initial conditions—variations which are *not* nomologically related to subsequent states of the Earth's atmosphere.

There are at least two ways to remedy the analysis to preserve a sense of nomicity. First, we might say that it is not a strict law, but a non-strict or probabilistic law that relates the case and effect. But this route does not look promising. For if it were only probabilistic laws that are needed in order to save the nomicity thesis, then the events in question could be subsumed by any number of descriptions such that if events of the same kind as the meteor impact, then events of the same kind as the extinction would occur fail to occur some given percentage of the time. If only probabilistic laws were necessary, then we would have a thesis that is true, but less interesting—presumably we are interested in finding out what features of the cause event were causally relevant to, or necessary for, the occurrence of the effect *in the particular case*, *i.e.*, one in which the effect *did* occur. Subsuming an event under a probabilistic law does not explain why, if the effect occurred, that it did, nor why, if the effect did not occur, why it did not. If, on the other hand, we were to take John Stuart Mill's route, and build in all of the relevant conditions such that when they are satisfied, then the effect always occurs, we would lose the intended implicit generality that motivates causal nomism: we would, in essence, simply be characterizing a unique causal process. It could still be true that the Chicxulub Meteor impact caused the dinosaur extinction even if there were no true law relating the events in question. It seems more likely that the event in question was unique, and that in order to provide a plausible explanation for the why the dinosaurs died out, we will need to construct a complicated, detailed model of that single event, the relevant features of which were unique in the history in the planet.

It is still possible, however, that if we decompose the complex event of the meteor impact and the complex event of the extinction into a number of simple linear causal chains of efficient cause-effect sequences, we might be able to preserve some interesting sense in which causality is nomological, and that singular causal relations are implicitly general. In other words, even if, for some (or all) complex causes and effects, there is no law under which the events can be subsumed, it may be that the simple events of which the complex events are composed exemplify strict causal laws.

G. E. M. Anscombe[24] objects to the arbitrary nature of the nomologicality postulate, maintaining that Davidson has offered no reason for thinking that it is true. I disagree: one of the theoretical motivations for accepting nomologicality is the possibility of employing the deductive-nomological model of explanation to the explanation of particular causings. This is an assumption which Davidson gives some indication of sharing. The explanation of some event *b* would require that there be some description of both *a* and *b* such that '*a* causes *b*' can be deduced from a causal

[24] Anscombe, "Causality and Determination," 104.

law of the appropriate syntax, together with the associated description of *a*. Evidently, a prior commitment to the inferential conception of scientific explanation drives Davidson to the view that there has to be *some* law relating *a* and *b* if *b* is caused—*i.e.*, if *it has a causal explanation*. D-N explanations need *some* law to cover the events. But in order to accept the nomologicality of causality even in this form, we need some reasons for thinking that the D-N model of explanation, which requires deductions from laws expressed as conditionals in some extensional language (together with a specification of initial conditions and auxiliary hypotheses) is the proper approach to causal explanation.

In recent developments in the theory of explanation, the D-N approach has come under considerable criticism. One aspect of that critique is its neglect of relative explanatory adequacy. If the D-N model is presupposed, then Davidson is correct to claim that "our justification for accepting a singular causal statement is that we have reason to believe an appropriate causal law exists, though we do not know what it is."[25] Yet, the justification we have for accepting the Chicxulub Meteor version of the Alvarez Hypothesis vis-à-vis the intrinsic gradualist hypothesis of dinosaur extinction is that the Chicxulub Meteor hypothesis explains the Iridium layer data and the gradualist hypothesis does not.

The argument for the nomological character of causality will, in all likelihood, depend upon adopting prior theoretical stands on the nature of causal laws and the nature of scientific explanation, and until such positions are articulated and defended, the exact sense in which causal nomism might be true will remain an open question.

§4. Causation and Counterfactuals

It is often assumed that counterfactuals are a necessary ingredient of an adequate general theory of causation. Assuming a regularity view of causation, it is easy to see why this a reasonable criterion. It is generally recognized that Hume's theory of causality has difficulties with the distinction between accidental and law-like regularities—both find their expression in universal conditional statements. But in order for the statements to express causal regularities and not just accidental constant conjunctions, the corresponding law-statements need to have a stronger modal force. One of the more commonly discussed ways of drawing the needed distinction is to say that universals of law imply or "sustain" counterfactuals, while mere "universals of fact" do not. The latter are accidental generalizations and do not have the same modal force.

Even the singularist will admit that causal claims have some counterfactual implications. I use Searle's example to illustrate the point.

> I am walking along when suddenly a man coming the other way accidentally stumbles into me, pushing me into the gutter. Now again, barring hallucinations and the like, I know without any further observation the answer to the question "What caused you to go into the gutter?" The man bumped into me and pushed me into the gutter. In this case, one wants to say "I know all of this because I *felt* myself being pushed into the gutter, and I *saw* the man doing it to me."[26]

Searle claims that the same evidence by which he knows that he was caused to go into the gutter also justifies his knowing the truth of the counterfactual claim if he had not, then and there, been bumped, *ceteris paribus*, he would not have ended up in the gutter. His knowledge that if, then and there, he had not been bumped, he would not have ended up in the gutter is not grounded in

---

[25] Davidson, "Causal Relations," 160.

[26] Searle, *Intentionality*, 119.

knowledge of any applicable laws nor even from the knowledge that there are applicable laws (granting that there are such covering laws).[27]

That is Searle's claim; is it correct, in general? Can Searle know the truth of a counterfactual by perception—by *actual* experience? It might have been the case that, unbeknownst to Searle, at the very same time as he felt and saw himself being bumped into the gutter, Searle was simultaneously tripping on a rock, which on a lonely path would have been sufficient under the circumstances to cause a fall into the gutter. If the effect was overdetermined in just this way, the effect might still have happened even if the ostensible "cause" did not. Searle would have to know that there was no other cause present that would have been sufficient, under the circumstances, to produce the effect. However, that does *not* imply that Searle would have to have known that there was some law or other instantiated when he was pushed into the gutter. The analysis of singular causal claims presupposed here appears to fit well with the one offered by C. J. Ducasse.

Recall from Chapter Two that Ducasse regards causality as the relationship between the factors of a special state-of-affairs called an "experiment." An experiment is a system in which any changes that occur in the system are in principle all perceptually detectable. Causation, or "the causal relation" is a triadic relationship between (1) a specific "state-of-affairs" in which only two changes (each either simple or complex) occur, (2) one of the aforementioned changes being A, beginning at t_1 and (3) the other, a change B, beginning after t_1. If such conditions hold, then

> The *cause* of an event *B* was an event *A* which, in the then existing circumstances, was *sufficient to* the occurrence of *B*; and therefore, conversely,
>
> The *effect* of an event *A* was an event *B* which, in the then existing circumstances, was *necessitated by* the occurrence of *A*.
>
> A condition of an event *B* was an event *A* which, in the then existing circumstances, was *necessary to* the occurrence of *B*.

An event that is causally sufficient for the occurrence *necessitates* the event, or, to use Ducasse's coinage, the cause is "etiologically sufficient" for the effect.

For Ducasse, an inquiry into the causes of an event is an inquiry into the immediately preceding changes in the proximity of the effect, regardless of how we might describe the changes in the system, or whether the two changes are nomologically related. If in a given otherwise stable system, two changes occur, Ducasse regards it as a certainty that the first was the cause of the second. Let us consider the definition that Ducasse offers for "*x* caused *y*,"[28] recalling that if the cause of an effect is the single change in the circumstances immediately preceding the effect, then the cause, in those circumstances, was causally *sufficient* for the effect:

D1 The change *x* occurred during a time and through a space termination at the instant *I* at the surface of *S*.

D2 The change *y* occurred during a time and through a space beginning at the instant *I* at the surface of S.

D3 No change other than *x* occurred during the time and through the space of *x* and no change other than *y* during the time and through the space of *y*.

---

[27] Ibid., 120.

[28] The definition is based on Ducasse, "Nature and Observability," 127. I have changed the letters in this example for consistency across the various authors' analysis of the meaning of the phrase—'y' is substituted for 'K' here, and 'x' for 'C.'

More briefly, the idea is that the cause of the particular change *y* was that particular change *x* as alone occurred in the immediate environment of *y* immediately before.[29]

Ducasse is of the view that a cause refers to the *single* change in the circumstances immediately before the occurrence of the effect in question which is sufficient to the effect. Ducasse emphasizes that for genuine causal necessity, the necessitating causal factor *x* must be the *only* immediately antecedent factor to change in the experimental system. His intuition here is at least understandable, for suppose that in S, three changes occurred, *x* beginning at t_1, *y* beginning at t_2, and *z* starting at t_3.[30] There are many possible causal states-of-affairs involving S:

- *x* caused *z*, and *y* had no causal relevance to *z* in S
- *y* caused *z*, and *x* had no causal relevance to *z* in S
- *x* caused *y* and *y* caused *z* in S
- *x* and *y* jointly caused *z*, but were individually insufficient to do so in S
- *x* and *y* were individually sufficient to cause for *z* in S
- *x* caused *z*, but *y* hastened or delayed it in S
- *y* caused *z*, but *x* hastened or delayed it in S

The multiplicity of possible arrangements of causal factors involved when the number of changes goes to three or higher makes it impossible to determine, without further experimentation in which either *x* or *y* is controlled, which of the two was the cause of the event *z*. In cases where only two changes are observed, no experimental repetition would be necessary to establish that *x* (or *y*) caused *z*, although the experiment might be repeated to make sure there were no mechanical problems with the set-up the first time, or that no human error (*e.g.*, in taking measurements) occurred.

Complications arise when we consider that such a definition of cause implies that *all* of the proximate changes in the vicinity of the effect are included in the cause, even those changes which intuitively we would regard as not causally relevant. Suppose, for example, that in S, changes *x* and *y* occurred simultaneously and satisfy D1-D3**,** but that *x* is not the kind of thing that is normally causally relevant to the occurrence of *z*-type events. Ducasse is not as worried about this as some commentators have suggested he ought to be.[31] Consider this example:

> At the instant a brick strikes a window pane, the pane is also struck perhaps by the air waves due to the song of a canary near by. Yet we usually would say that the cause of the breakage was the impact of the brick, and that the impact of the air waves, although it was part of the prior total change in the give state of affairs was no part of the cause.[32]

If, as we are accustomed to do, we refer to the brick's striking the window as causing its subsequent breakage (*i.e.*, it's *just* the brick, not the canary song, that counts here) then we are employing a notion of cause that departs from that encapsulated in his definition, since we are extracting only one (albeit dominant) aspect of the *total* cause for inclusion. The intuitive basis for this is our knowledge that the brick's motion would have been sufficient, even if the canary had remained

---

[29] Ibid.

[30] Ibid., 151.

[31] Mackie, *Cement*, 134ff.

[32] C. J. Ducasse, *Nature, Mind and Death* (La Salle: Open Court, 1951), 123.

silent, to break the window. The same could not be said for the efficacy of the canary's song. Ducasse's definition of cause as single proximate change cannot be satisfactory if we still need some basis on which to distinguish the causally relevant aspects of the change from the causally irrelevant. The definition is too broad, it would seem.

According to Ducasse, the objection misses the mark, because it equivocates on the type-token distinction. If we start with the question "What caused *y*?" we can presumably mean one of two things:

(1) What token event *x* caused the token event *y* (the change in the subject S) such that *x* was the only change that satisfies conditions D1-D3?

(2) In virtue of exemplifying what property-type(s) *A* did the token event *x* cause the event *y* to exemplify property-type(s) *B*?

According to Ducasse, (2) presumes an answer to (1), but goes on to ask what aspects or features of *x* were causally relevant to the occurrence of its effect—*what part of* x *is left if we subtract from* x *those properties of, or parts of* x *which were unnecessary to the occurrence of* y, *leaving just those properties or parts of* x *which were sufficient for some part of* y*?* In the canary-brick example, the cause of the window's breaking is *all* of whatever changed in the causal context of S immediately before *y*. In other words, the cause of the window's breaking includes "the song of the canary, the precise mass shape, position and speed of the brick" and whatever else was in any way causally relevant (*i.e.*, necessary) to *any* aspect of the effect, or causally relevant to the effect, under any description of it. The canary's singing was not causally relevant to the breaking of the window *qua* breaking—it was not necessary to the occurrence of the breakage. Nonetheless the bird song had effects, "and these too were parts, but *other* parts, of the whole later event."[33] Insofar as we are interested in acquiring practical causal knowledge for prediction and manipulation, we are interested in finding out what *types* of causes under what conditions are sufficient for certain *types* of effects. Ducasse acknowledges this. We are often interested in the specific properties of events to which we can assign causal relevance in our explanations of event types. But we must not be mislead by the superficial features of linguistic usage, in which type/token confusion is often implicit, to think that such usage reflects any interesting ontological facts about causality. Causes and effects are particulars, not types.

Mackie criticizes Ducasse's definition on the grounds that although "every cause is sufficient in the circumstances for its effect … so are many non-causes for events which are not their effects."[34] In effect, Ducasse's definition amounts to a (very) weak sense of sufficiency. Mackie's alternative is a "strong" or "counterfactual supporting sense" of causal sufficiency, with the implication that if, in those very circumstances, *y* were not going to occur, then neither would *x* have been going to occur. Those aspects of the cause that were necessary as well as sufficient need to be included in the specification of the cause. According to Mackie, the question "What was necessary in the circumstances for *y*" is equivalent to "What occurred such that if it had not occurred, *y* would not," and so it is the necessity of causes for their effects that gives the required counterfactual import of singular causal claims:

> … the element of counterfactual conditionality, of the cause's being necessary in the circumstances for the effect, is part of our ordinary notion of a singular causal sequence,

[33] Ibid., 124.

[34] Mackie, *Cement*, 39.

> and it cannot be brushed aside as an intrusion of the general into the singular, as involving a concealed comparison of this individual sequence with others.[35]

If a stronger, counterfactual-supporting sense of causality is needed, the appropriate formula is given by the specification of the INUS condition for an event *y*.[36] According to Mackie, *x* caused *y* entails that *x* is an insufficient but necessary part of a condition which is itself unnecessary by sufficient for *y*.[37] More formally, the basic idea is this:

X caused $y \equiv_{df}$

> *X* and *y* occurred; and *X* was that set of conditions, among the totality of those that actually occurred, each condition x_i of which was necessary for the occurrence of *y*, and the entire set $X=\{x_1,x_2, \ldots x_n\}$ was sufficient for the occurrence of *y*, and *y* followed upon *X*.

The problem with Mackie's argument for the counterfactual import of singular causings is that it begs the question against Ducasse. Ducasse rejects the idea that the meaning of a singular causal statement—a statement of what *did* happen—can be analysed in terms of what did *not* happen. The imaginary counterinstances required to sustain Mackie's alternative analysis are ruled out on Ducasse's account. Mackie merely contradicts Ducasse's point about the admissibility of imaginary counterinstances, without arguing against it. For Ducasse, to say that "the song of the canary was unnecessary … is to say only that in *another* case, otherwise similar, where the song did not occur, an effect of the *same sort* viz. breaking, nevertheless did occur."[38] What is necessary or unnecessary can therefore only be established with reference to generalizations over facts about similar cases, and that is why Ducasse argues that counterfactual import is no part of particular instances of causality. In fact, Mackie admits as much:

> The singular causal statement says that without A, on this occasion … P *would not* have occurred; this is very often supported by the observation that without A, on some other similar occasion … P *did* not occur. … this 'did not' supports the 'would not' because it is assumed that there is *some* underlying regularity of behaviour …[39]

I agree with Mackie that it is only in an "unsophisticated"[40] sense, where we imagine what did not happen in the place of what did, that causality is implicitly general. To be fully consistent, I submit that we need to disabuse ourselves of the intuition (if we have it) that singular causal relations necessarily have counterfactual import.

There is one other argument that could be used to support this conclusion. Suppose one were to argue as follows:

1. If causal relations, as such, are in principle detectable in perception, then we would have to be able, in principle, to detect the aspect of the relation that is essential to its *being* a causal relation.
2. What is essential to being a causal relation is having counterfactual import.
3. Counterfactual import is not detectable in perception.

---

[35] Ibid., 140.

[36] Mackie hedges his position considerably, only maintaining that this is *often* what *we* (speakers of English) mean when we say "*x* caused *y*."

[37] J. L. Mackie, "Causes and Conditions," in *Causation*, ed. Ernest Sosa and Michael Tooley (Oxford: Oxford University Press, 1993), 35.

[38] Curt J Ducasse, *Truth, Knowledge and Causation.* London: Routledge & Kegan Paul, 1968), 12-3.

[39] Mackie, *Cement*, 79.

[40] Ibid.,

∴∴ 4. Therefore, causality is not detectable in perception.

Now, we agree with Mackie and Ducasse that "causation can be *detected* in a single sequence,"[41] so the conclusion (4) is false. But the premises (1) and (3) are sufficiently likely to be true that we would be wise to reject (2).

There are some signs that an anti-counterfactuals "backlash" has been gaining some adherents of late: witness Wesley Salmon's recent attempt to present a counterfactuals-free analysis of causality.[42] The theoretical motivation for wanting a counterfactuals-free analysis of causality is part of a related, but somewhat larger issue: the goal of presenting

> a "process theory" of causality that could resolve the fundamental problems raised by Hume regarding causal connections. The main point is that causal processes, as characterized by this [*i.e.* Salmon's] theory, constitute precisely the objective physical causal connections which Hume sought in vain.[43]

Causal processes were used in Salmon's theory as part of the explanation of the transmission of causal influence—to explain how an action at one point could result in the action of something else at a subsequent time, some distance away. According to the theory of special relativity, the transmission of causal influence such as exemplified by the propagation of electromagnetic waves through a vacuum, have a maximum velocity of *c*, the speed of light. However, there are processes that *could* travel at faster than light speeds; there is no real limit to how fast such processes can go. These superluminal processes cannot, however, be used to transmit signals—they are "pseudo-processes."

How do we distinguish causal processes from these pseudo-processes? According to Salmon, the causal processes (such as beams of light) are capable of transmitting marks, or information, from one place to another, while pseudo-processes are not. In his formal definition of mark transmission, Salmon employed a counterfactual specification of the nature of causal processes. As Kitcher noted, "the counterfactuals are the heart of the theory."[44] Arguably, any analysis of natural causality that requires counterfactuals presupposes semantic causal nomism, and the implicit generality of causality. This admission stands, as Kitcher argued, as a serious difficulty for singularist theories of causality and for the ontic conception of causal explanation.

Salmon's recent successful expulsion of references to counterfactuals in the analysis of the basic ontology of causation is, I think, theoretically highly significant, and illustrates that quite different theoretical approaches can end up converging on a singularist account of causation. There is yet a third theoretical approach to causation that *prima facie* converges upon the singularist conclusion—the account developed in recent years by Micheal Tooley.

3.3 Singularism and Non-Reductivism

Michael Tooley[45] has recently defended a version of singularism as part of an attempt to

---

[41] Mackie, *Cement*, 135, n. 31.

[42] Wesley C. Salmon, "Causality Without Counterfactuals," *Philosophy of Science* 61, no. 2 (1994): 297-312.

[43] Salmon, Causal Structure of the World, ch. 5.

[44] Kitcher, "Explanatory Unification," 472.

[45] Michael Tooley, "Laws and Causal Relations," in *Causation and Causal Theories*, Midwest Studies in Philosophy, vol. IX, ed. Peter A. French, Theodore E. Uehling, Jr., and Howard K. Wettstein (Minneapolis: University of Minnesota Press, 1984), 93-112; idem, *Causation: A Realist Approach* (Oxford: Clarendon Press, 1987); idem, "The Nature of Causation: A Singularist Account," *Canadian Journal of Philosophy* Supplementary Volume 16 (1990): 271-322; idem, "Causation: Reductionism vs. Realism," in Ernest Sosa and Michael Tooley,

articulate a realist theory of causation. Based on the terminology he uses to characterize his conclusions, one might expect to find similarities between his goals and mine, and that is the case. However, deep differences in methodology and ontology between us lead us to divergent conclusions on a range of issues. My assessment and critique of Tooley's argument for singularism will call for the explicit treatment of some side-issues that will highlight how I believe that a realist theory of causality ought to be defended, and why an aspiring realist is better served by my approach than by Tooley's.

Much of Tooley's project is devoted to assessing the relative merits of contrary positions on a number of related issues:

- Supervenience *vs.* Singularism
- Reductionist *vs.* Realist Singularism
- Realism *vs.* Reductionism of causal laws
- Realism *vs.* Reductionism of causal relations

§1. From Supervenience to the "Intermediate View"

The position characterized as the "Thesis of the Humean Supervenience of Causal Relations" is a conjunction of these two theses:

S1 Causal relations presuppose corresponding causal laws

S2 Causal relations are logically supervenient upon causal laws plus the non-causal properties of, and relations between, events.[46]

The idea here is that being a specifically *causal* relation is conceptually dependent on there being some causal law or other (known or not) which the causal relation exemplifies. There is a semantic dependency of "causality" upon "nomicity." Elsewhere, Tooley offers a somewhat more perspicuous and more general formulation of **S2**:

S2' The truth values of singular causal statements are logically determined by the truth values of statements of causal laws together with the truth values of noncausal statements about particulars.[47]

Together, these two claims are very similar to the "semantic" component of causal nomism in §3.1, which are both entailed by the formal statement of $\mathbf{CN}_S$ given in §3.2. Singularism is, of course **~S1 & ~S2.** There is, however, a third possibility, according to Tooley, because the conjunction **S1 & ~S2** is logically possible. That is, an "intermediate view" that retains the Humean idea that causal relations presuppose causal laws, but drops **S2.**

His first step is to distance himself from the type of argument I have characterized as the "detectability argument" for causal singularism. Without naming the sources, Tooley indicates four ways in which the thesis of the "immediate knowledge" of causal relations has been intended:

(1) Causal relations can be given in immediate experience, in the strong sense of actually being part an experience itself.

(2) Even if causal relations are not given in immediate experience, one can …have noninferential knowledge that states-of-affairs are causally related.

(3) Causal relations are … observable in many cases.

---

eds. *Causation* (Oxford: Oxford University Press, 1993), 172-91.

[46] Tooley, "Nature of Causation," 274.

[47] Tooley, "Laws and Causal Relations," 93.

(4) There are situations where observation of a single case can provide one with grounds for believing that two events are causally related, and that it can do so even in the absence of any prior knowledge of causal laws.

Clearly, (1), (3), and (4) are equivalent, all slightly different formulations of the detectability thesis I advanced in Chapter Two. Recall that the claim there was made that we can detect or perceive efficacious agency, and so those forms of detection deliver immediate non-inferential causal knowledge.

Tooley's only argument against the detectability thesis—the immediate experience of causality—is wide open to refutation on a number of points. If this is the best refutation that Tooley, a master of creative counterexample, can come up with, that makes me all the more certain that the detectability thesis is true. Tooley's argument is an easily recognizable variant of "brain-in-a-vat" skepticism; the idea is that for any process of detection that depends on the operation of a causal process (*e.g.*, perception, sensation), if we were sufficiently clever, we could artificially induce a phenomenologically identical experiential state as would be had by the operation of the normal causal process *in the absence* of such processes. Tooley's argument[48] is basically this:

T1 If a relation between *a* and *b* is immediately given in experience, then it is part of the content of the experience .

T2 If something is part of the content of an experience E, then it is also part of any experience E* that is qualitatively identical to E.

T3 Given any experience E of a relation between *a* and *b*, it is logically possible that appropriate, direct electrical simulation of the brain might produce an experience, E*, which was qualitatively indistinguishable from E, but in which the experiential contents *a* and *b* were not causally related, being the independent products of a common cause.

∴ For any immediately given experience of a causal relation between *a* and *b*, it is logically possible that *a* and *b* are not causally related, being the independent products of a common cause—an appropriate direct electrical stimulation of the brain.

There are several things wrong with this argument. First, there is equivocation on the term "experience." In **T1**, "experience" is employed relationally, as being identical with the sum of the intentional *contents* experienced. In **T2**, "experience" is used non-relationally, to refer to the phenomenal aspects of the experiential state, disregarding the intrinsic intentionality of such states. Moreover, **T2** is false. My perception of an egg's shape and colour is part of the content of an experience of an egg today, but that does not mean that the very same egg is part of the content of my experience whenever I have the qualitatively identical experience of examining another egg. How could it be, when each egg ends up getting cooked and eaten?! Thirdly, **T3** is an arbitrary assertion. There is no neurophysiological evidence to support such a fantastic claim, and thus no grounds for thinking that the imagined state of affairs is in fact possible. The fact that such a scenario could be imagined is even dubious, for in order to imagine it, one would have to have sufficient knowledge of neurophysiology and psychology to get a clear and distinct mental image of the chain of events involved, from the attaching of electrodes, to the experimenter's getting the subject to report that she has had an experience of causality. But the more relevant knowledge one possess, the more utterly fantastic assumptions one would have to swallow with no empirical evidence to wash them down. Furthermore, even if it were psychologically possible for a knowledgeable neurophysiologist to imagine how to perform the "appropriate" stimulation, that is not

[48] Tooley, "Nature of Causation," 284.

even "in-principle" support for the suggestion that such an appropriate stimulation could be externally delivered. Fourthly, **T3** is self-refuting: I would have to already know the difference *in experience* between the naturally caused experience and the artificially induced experience in order to even formulate the distinction between them. That presupposes that I have had prior epistemic access to some naturally caused experiences, *i.e., that* I *had* a naturally caused experience (non-artificially induced) and could recognize it as such.

So much for the cogency of Tooley's attack on the detectability of causation. Indeed, in the three years after his argument was published, he backpeddled considerably. In a more recent article, he concedes that "if [perceptual] direct realism is correct, then I think it can be plausibly held that one has non-inferential knowledge of causal states-of-affairs."[49] The only way that the argument just given could be motivated is if "a satisfactory account of non-inferential knowledge requires … [that] what properties one is directly acquainted with is logically supervenient upon the phenomenological content of one's experience."[50] But that would make the position incoherent. If the properties directly perceived are logically supervenient upon the phenomenological content and not the objective content of the experience, it is a phenomenalist view of non-inferential knowledge that is being advanced, which is incompatible with direct realism.

Tooley's worry that there is "no epistemologically neutral way of showing either that one can or that one cannot, have non-inferential knowledge of causal relations between events" is based on a confusion, killing any valid motivation for pursuing the view that causal relations are theoretical relations. What is even more puzzling is that there was no good motivation for his attack on the detectability thesis in the first place. For if Tooley thinks that the detectability thesis *does not* imply unanalyzability of causality, then there is no reason why he should be attacking the thesis, or Anscombe. All he needed to do was to reject the eliminativist route and propose a positive singularist analysis of causality.[51]

Let us assume for the sake of argument, that Tooley is right, and that detectability-based or "reductionistic" singularism is a failure. That means we have to look to non-reductionistic approaches to singularism.[52] That means, first, that we need *new* arguments against the Humean supervenience view and its reductivist implications, *i.e.,* arguments that replace the detectability arguments for singularism with other, more promising ones. Tooley's strategy is to mobilize arguments against "the supervenience view," which, if successful, would leave the "intermediate" and singularist theories as the only live options. He then will offer additional considerations for preferring singularism to the intermediate view.

Over the years, Tooley has developed at least six different arguments against the supervenience view. Granting that supervenience implies reductionism, they are arguments against the view that causal relations are *reducible* to non-causal states-of-affairs. The upshot of the arguments is that one can imagine possible worlds where the causal relations between events are logically underde-

---

[49] Tooley, "Reductionism vs. Realism," 191.

[50] Ibid.

[51] Tooley finds a legitimate target in Ducasse's analysis of causality. But the objections he raises are spurious. (Tooley, "Nature of Causation," 287-88) Tooley thinks that Ducasse's account rules out the possibility of (i) uncaused events, (ii) telepathic communication, and (iii) discontinuous or "gappy" causal processes. Perhaps Ducasse's account has these three implications. The more important point is that nothing Tooley says would persuade the close reasoner that these implications count against Ducasse's definition.

[52] The approach Tooley has in mind is one in which a causal relation is any relation between states-of-affairs which satisfies a set of postulates governing probabilistic causal laws.

termined by the conjunction of their associated causal laws plus the non-causal properties of, and relations between, events. The arguments themselves feature such bizarre entities as "rotationally symmetric universes," "inverted universes," "uncaused events," and "indeterministic laws." They prove nothing about the world except the fertility of Professor Tooley's imagination. But that is just fine, because Tooley is not even *interested* in talking about the world. His goal is

> to set out an analysis of the *concept* of causation, and not merely to offer an account that is true of causation as it is in the actual world. The theory must be true of causation in all possible worlds. So none of the statements in the theory can be merely contingently true.[53]

This criterion of adequacy for causal theories leads to an unsound methodological orientation. Here is why. Suppose that we arrive at a serviceable concept of causality, reducible to some immediately detectable aspects of experience, and formed by the process of abstraction outlined in §2.5. Now suppose our possible döppelgangers in the other possible worlds do the same thing, and come up with their own concepts of causality. If Tooley is right, then *our* concept of causality must be defined by the essential trans-world similarities between causality as it exists in our world, and as it exists in the other possible ones. That gives us this dilemma: either the characterization of causality which is valid for our world (a "contingently true" account) is the same as the characterization of causality that is valid across possible worlds (the "analytically true" account) or it is not. If it is, then the apparatus of possible worlds is theoretically dispensable in this context, and we can focus on talking about this world. If it is not, then *if* we prefer the trans-world characterization, and take it as superseding whatever contingently true account we have for this world, then we will be giving epistemic priority to the possible over the actual. If we desire a theory of causation that fosters empirical science, that can lead to improvements in the design of, and interpretation of experiments, then we need an account of causation that is true of the actual world, not other possible worlds.

If the characterization of causality which is valid for our world is the same as the characterization of causality that is valid across possible worlds, then the possible-worlds formulation is redundant. If it is not, and we opt for the trans-world characterization, we are led to a view of causation detached from the conceptual and methodological needs of science, and therefore to a concept of causation that violates one of the basic conditions of conceptual validity: cognitive utility (see Chapter Four).

§2. From the "Intermediate View" to Singularism

So far we have established that Tooley's arguments against detectability-based causal singularism are without logical standing, and his aspirations to formulate a non-reductive singularism unmotivated. We have seen that his "replacement" arguments against the Humean supervenience view are based on faulty methodological assumptions. Even if none of these problems had arisen, it would still be possible that the position characterized as the "intermediate view" might still be right, so Tooley needs to advance an argument for the conclusion that **S1** ought to be rejected (as well as **S2**). If that argument goes through, then it is "logically possible for there to be causally related events that do not fall under any law." In other words, "events can be casually related even in worlds where there are no causal laws."[54] Whether *our* world is such a world is of

---

[53] Tooley, "Nature of Causation," 292.

[54] Ibid., 322.

crucial importance, but Tooley's methodology is incapable of getting an answer to this question. At any rate, if such worlds are possible, then our world *could* be one. So it is still worthwhile to examine the argument to see if worlds in which there are causal relations but no causal laws is possible.

The argument invokes considerations of simplicity. The basic idea is that the ontological commitments of singularism are more austere than those of the intermediate view, and so the former view is more plausible than the latter, by implicit appeal to Occam's razor.

> The singularist approach can … account for everything while postulating only two special sorts of facts. For although it cannot reduce causal laws to causal relations between states of affairs, it can analyse the concept of a causal law in terms of the concept of a law … together with the concept of causal relations.[55]

On the other hand,

> an intermediate account needs to postulate three special sorts of facts: those corresponding to non-causal laws, those corresponding to causal laws, and those corresponding to causal relations between states of affairs.[56]

Since simpler is better, the singularist view wins.

In brief, how does Tooley get from the Humean view to the singularist view? First, he gives an invalid argument against the detectability thesis, which he attacks despite the fact that it does not imply the unanalysability of causality, and despite the fact that there are good arguments in favour of thinking that causal relations in non-scientific contexts are detectable. Second, he assumes, despite lack of motivation for it, that causality is best understood as a theoretical relation. Third, he offers several bizarre thought experiments invoking imaginary "logical possibilities" to refute the supervenience thesis. Fourth, as we have just seen, he invokes Occam's razor to make the last step in the argument for singularism. Now this is somewhat less controversial as a methodological posit than others in Tooley's bag of tricks. But it has the status as a "rule of thumb," and it is inappropriate for Tooley to place such heavy theoretical demands on such an inferentially weak principle.

My approach, on the other hand, gives direct evidence in favour of singularism on the basis of the detectability thesis, combined with an abstractionist approach to the ontology of causation. Notice that since Tooley is, at best, skeptical about the arguments for the direct detectability of causality, he cannot endorse the idea that singular causal claims can be immediately justified by the detection of causality. In other words, he accepts the idea that the concept of causality presupposes the concept of a law of nature, as the Humeans do. He accepts the burden of proof, and takes the stance that the supervenience view must be refuted. But this stance is mistaken. Since we have non-inferential knowledge that some relations are known to be causal (*e.g.*, the relation between a decapitation and a death, or an intentional decision and the consequent action), such knowledge depends upon no causal laws, known or otherwise. There is no evidence available in experience that supports a causal nomist view. The onus of proof properly rests upon the Humean, who must show that causal claims worthy of the title "causal" presuppose some law or other. If no sound argument is forthcoming, it is at best an open question, and at worst, an arbitrary postulate.

My approach is therefore considerably more powerful than Tooley's, both inferentially and in terms of its philosophical fruitfulness. It is much more congenial to causal realism than Tooley's

---

[55] Ibid., 277.

[56] Ibid., 278.

own, besides. There is one last matter connected with Professor Tooley's article that requires special attention, however.

Tooley wants to argue for the conclusion that "the concept of causation is parasitic upon the concept of a law of nature." Nonetheless, he thinks that "this sort of conceptual dependence need not entail any ontological dependence. Events could still be causally related without falling under any relevant law."[57] How, on the one hand, can Tooley be justified in denying **S1,** that is, saying that causal relations *do not* presuppose corresponding causal laws, and at the same time maintain that

> even if one adopts a singularist conception of causation, one is never justified in believing that two events are causally connected unless one is also justified in believing that there is some causal law of which the relation in question is an instance.[58]

In other words, how is it that these three claims can all be true?

1. Causal relations do not presuppose corresponding causal laws.
2. Events can be causally related without falling under any relevant law.
3. One is never justified in believing that two events are causally connected unless one is justified in believing that there is some causal law of which the relation in question is an instance.

To maintain that all three are true is to endorse the epistemic component of causal nomism while rejecting the semantic component. Now, it might occur to you to ask this question: if one is never justified in believing in anomic singular causal relations, what other grounds are there for thinking that this is so besides the Humean view that singular causal claims are implicitly general? Is it not the case that the denial of semantic nomism *implies* the rejection of epistemic nomism?

If Tooley's argument for singularism were the only available option, causal nomism would seem to be preferable, given the rather bizarre flights of imagination necessary for Tooley to come up with counterexamples to nomism that he finds convincing. However, in light of the causality-detectability arguments for causal singularism that establish its *prima facie* plausibility, and the absence of arguments from Humeans in favour of regarding semantic causal nomism as true, singularism is the more defensible position of the two.

3.4 The Ontology of Causal Relata

What *kinds* of things are causes?

According to the most common view, *events* are what enter into causal relations. There are compelling reasons, however, for thinking that this view is mistaken, and that entities—*powerful particulars*—are causes. To add another dimension of confusion to the problem of causal relata, it is often the case that we will speak of *facts* as being causes, such as if we were to say "the fact that its temperature is 1200°C is causing the metal to glow," or "the fact that the baby was hungry caused it to cry earlier." So there seems to be at least three ways in which we can refer to causes: as agents (or perhaps more generally as objects), as facts, or as events. Fortunately (but somewhat puzzlingly) we do not have this problem with effects. It is generally conceded that they are events, or actions.[59]

Before we can tackle the ontological problem, there are some terminological issues that need

---

[57] Ibid., 293.

[58] Ibid., 321.

[59] This distinction is unproblematic because actions are a subset of events, those events involving changes in an entity. For example, "the end of winter" is an event but not an action since there is no specific particular which is undergoing a change.

to be dealt with.

§1. *A Taxonomy of Causal Claims*[60]

The most basic causal claim is one of the form "*x* caused *y*" or "*x* is causing *y*." Such expressions are commonly called "singular causal claims." The causal relata named by "*x*" and "*y*" refer to non-repeatable, dated, particulars. Hence, a singular causal claim stated in the form "*x* caused *y*" is elliptical for "*x* caused *y* then and there." On the other hand, *x* and *y* in judgments of the form "*x* causes *y*," are not properly regarded as particulars, but as types, such as "viruses cause diseases." The relata identified in this second kind of causal judgment are natural kind categories, not particulars. There are basically two conceptually distinct kinds of judgments that may be stated grammatically the same way.

Some statements of the form "*x* causes *y*" are what James Woodward calls "causal role claims." The claim "smoking causes lung cancer" is one of this type. In general terms, these are claims that "some causal factor or process is playing the role of producing some effect (commonly an effect whose magnitude or rate or occurrence can be given some … quantitative description) in some population."[61]A bit more precisely, these are claims about some causal factor *C* equivalent in meaning to "the frequency / quantity of effect *E* in population *P* is different when *C* is present than it is if *C* is absent, all else remaining the same." The idea here is that the presence of the causal factor *C makes a difference* to *E* in *P*. The causal relation between factors in the population is not necessarily instantiated in each member of the population. It is perfectly consistent with the claim "smoking causes lung cancer" that only 5% of smokers in some population ever get lung cancer. The claim implies that in an otherwise identical population with fewer smokers, the incidence of lung cancer would be proportionally lower, without implying anything about the relative rates of risk in *particular* individuals. Quantitative methods involving randomized clinical trials, linear causal modeling, and other techniques are widely used in the biomedical and social sciences to test causal role hypotheses.

The most abstract causal propositions are those in which "*x* causes *y*," means "*x*s are the type of thing that have the potential to make *y*-type events occur." I refer to expressions like this as "capacity ascriptions." Capacity ascriptions attribute to particulars the power or the latent potential to bring about some kind of state of affairs in virtue of being the kinds of things they are. For such claims as "smoking causes lung cancer," it is ambiguous whether the claim is expressing a causal role claim about the increased cancer risk due to smoking, or whether it is expressing a statement about the causal capacities of the inhalants. If it is the former, then the claim is implicitly about a particular population. If it is the latter, then it is a claim about inhalants, not a claim about relative risk. If an entity has a capacity, it has it regardless of where it is—since the capacities of things are associated with their persistent features, an entity takes its capacities with it wherever it goes, so long as it remains the kind of entity it is essentially is—*i.e.*, until it is destroyed. On the contrary, the causal role of a factor may vary considerably across populations.

§2. *Entities as Causes*

According to the theory of causality-detection as the awareness of volitional efficacy, *agents*—

---

[60] In this section, I follow the taxonomy of causal claims suggested by James Woodward, "Capacities and Invariance," in *Philosophical Problems of the Internal and External Worlds*, ed. John Earman, Allen I. Janis, Gerald J. Massey and Nicholas Rescher (Pittsburgh: University of Pittsburgh Press, 1993), 283-328.

[61] Ibid., 285.

people—were regarded as the causes of their actions. Persons are, on this view, self-determining entities in the sense that the person has the capacity to cause actions to occur autonomously. If persons are agents in just that sense, they are not "antecedent sufficient conditions" of the actions that they cause to happen. Nor, obviously, are persons events. The differences between the event conception of causes and the agent conception of causes might be sufficiently different as to warrant different concepts to designate their mode of relationship to the effects that they determine. As Richard Taylor puts it,

> this conception of the causation of events by things that are not events is, in fact, so different from the usual philosophical conception of a cause that it should not even bear the same name, for "being a cause" ordinarily just means "being an antecedent sufficient condition or set of conditions." Instead, then, of speaking of agents as *causing* their own actions, it would perhaps be better to use another word entirely, and say, for instance, that they *originate* them, *initiate* them, or simply that they *perform* them.[62]

The enmity towards agent causation reflects a certain view of the nature of things—that things act and interact in the manner of efficient causation, that circumstances external to the entities involved are the primary determinants of their behavior. Mario Bunge characterizes this view as the doctrine that

> … a thing can acquire its determinations solely by the action of efficient causes external to it; … properties are assumed to cling *ab extrinseco* to a passive substance whose sole peculiar property is the capacity of receiving properties, being otherwise immutable, independent, and simple …[63]

Agent causation, on the other hand, leaves room for circumstances internal to entities having causal roles. In the mode of agent causation, entities' actions are to an extent *self-determined*, and therefore, to the extent that they are self-determined, their actions cannot be causally explained without reference to their "natures."[64]An entity acting—that is, *acting causally*—is an entity which is either (a) causing action to occur in itself, by means of some process initiated by, and intrinsic to, the entity itself, which cannot be reduced to the efficient causal influence of other external bodies or forces, (b) causing action to occur in others, by means of initiating a causal process whereby causal influence is transmitted by means of some mechanism, or (c) reacting to the causal influence of the action of other entities. Many actions of entities involve reciprocal causation in which modes (b) and (c) are occurring simultaneously; causal interaction is a type of reciprocal causation.

Agency is generally understood to be attributable only to sentient beings, it being regarded as anthropomorphic to attribute agency to simple material objects. For this reason, it is appropriate to regard agent-causation as one species of "entity causation," in which the intrinsic nature of the entity involved in an action is the dominant or salient causal component. There are, of course, other modes of causation in which the entities acting are mainly reacting passively to external causal influences—planets moving in their orbits under the gravitational pull of the sun, moons orbiting planets, and the interactions of billiard balls being the paradigmatic cases.

It has been suggested that "the concept of 'agent' causation can be understood as a distinct species (from 'event' causation) of the primitive idea [of] 'causal production,' underlying realist or

---

[62] Richard Taylor, *Metaphysics*, 3d ed. (Englewood Cliffs, NJ: Prentice Hall, 1983), 49.

[63] Mario Bunge, *Causality* (New York: Meridian, 1963), 198-99.

[64] See Chapter Four, below.

non-Humean conceptions of event causation."[65] I propose that this picture be modified according to a suggestion by Harré and Madden. They suggest that neither pure autonomy nor pure passivity ever adequately characterizes the mode of action of entities, but that

> a wide range of different proportions of responsibility is assigned on the one hand to intrinsic factors such as the constitution or nature of things, materials, and persons, and processes internal to them, and on the other to extrinsic factors such as the stimuli to which they respond. References to intrinsic states emphasizes agency, and to extrinsic circumstances emphasizes passivity.[66]

While particular actions are causally determined or necessitated by the specific circumstances of their action, this does not imply a doctrine of causal determinism. The identification of internal causal processes that contribute (indeterministically) to their effects is clearly mandated by the fact of personal volitional agency. As well, the singularist denial that causality is essentially nomological in character suggests that there are no strict, true, deterministic causal laws. We need to recognize several different components of causal determination, including, but not limited to, self-determination and efficient causal determination, which refer to the intrinsic and extrinsic components of a resultant action, respectively.

Given the solution to the "extension problem" offered in Chapter Two, there is nothing to invalidate conceptualizing all entities as having a nature upon which supervenes the potential to act, or the liability (or disposition) to react in specific ways given the appropriate circumstances and conditions, both external and external. This reconceptualization of causality in which entities' actions are regarded as being determined in part by their natures does indeed entail the need to elaborate a theory of what these natures are, and what the concepts of capacities and liabilities contribute to our understanding of motion and change.

As for the question of the relata of singular causal judgments, the implications of the entity-causation paradigm is simply this: that it is no less correct to say that "an entity caused the event" (*i.e.,* an action) than saying "an event caused the event." Strictly speaking, *neither* is true. Rather than conceiving of causal instances as relations between token events, we can conceive them more generally as actions in which some capacity or liability for action has become actualized by means of some process involving the contribution of the intrinsic nature of a thing as well as the contribution of a thing's "circumambient" conditions, to use Harré's phrase. In short, causal instances are instances of *power actualization processes.*

Interestingly, this view has been rejected by C. J. Ducasse, with whom on other issues pertaining to causality I have been in agreement. He maintains that when competent speakers of the English language employ the expression "is the cause of," they are talking only about events. Ducasse's otherwise deviant approach to causal questions is strictly Humean here:

> When we say that the carpenter made the table, that is, caused it to exist, what we really mean is that a certain set of the carpenter's *movements* caused in the lumber a *change of form* from that of a heap of boards to that of a table. … It thus appears that "objects" in the sense of things or substances are not the kind of logical entities that strictly can, meaningfully, be referred to as causes and effects.[67]

While Ducasse recognizes that objects (*i.e.*, entities) cannot be eliminated from our ontology, he does

---

[65] Timothy O'Connor, "Agent Causation," in *Agents, Causes, and Events: Essays on Indeterminism and Free Will*, ed. T. O'Connor (Oxford: Oxford University Press, 1995), 174.

[66] Harré and Madden, *Causal Powers*, 83.

[67] Ducasse, Types of Necessity, 52.

insist on an ontological separation of entities and their actions, reserving only a place for the latter in causal relations:

> It is things and not their changes which are agents and patients, that is, which do and undergo; on the other hand, it is the changes or the states of things, and not the things, which are causes and effects.[68]

This subtle separation gets Ducasse into considerable trouble in his revealing exchange with Sterling Lamprecht. The center of the debate involves the interpretation of what happened to a car that ran out of gasoline and stopped moving. The question is: what caused the car to stop?

Lamprecht's answer is: friction between the road and the tires.[69]

Ducasse's answer is: the disappearance of the gasoline.[70]

Ducasse thinks that Lamprecht's answer has to be wrong because it is only *changes* that have effects. The friction between the road and the tires counts as a "condition" because it is present throughout. The friction is causally relevant, as a part of the conditions in which the cause operates, but it cannot be counted as *the* cause. Likewise, Ducasse thinks that a *change* in friction upon the tires, such as would be caused by applying the brake, would be a cause of the car's stopping if it occurred. He uses a symmetry argument to drive home the point:

> Professor Lamprecht would doubtless admit that *increase* in the quantity of gasoline admitted to the cylinders does … *cause* an increase in the speed of the car. Would he then refuse to admit that, under the same conditions, a *decrease* in the quantity of gasoline admitted likewise *causes* a decrease in the car's speed?[71]

It follows from this line of reasoning that a decrease to zero injected gasoline would cause a stoppage of the car.

Lamprecht regards Ducasse's analysis of the situation as "harmless in practical affairs" but "technically quite inexact."[72] The metaphysical point upon which Lamprecht insists is suggestive: *only existences can be causes*. What does Lamprecht mean by "existences?" To motivate his answer, we ought to first say in what respect Ducasse's analysis is said to fail, and then consider the problem raised in more abstract terms. Here is how Lamprecht dismisses Ducasse's explanation of the stopping of the car with reference to the disappearance of the gasoline:

> This disappearance of gasoline is a fact about the concrete existences we are describing. It is not … one of the existences described. We cannot legitimately say that the disappearance of any thing exists. And if it does not exist, it can hardly be causally operative.[73]

This implies that Lamprecht would reject the symmetry premise implicit in Ducasse's argument. It does not follow from the fact that only existences can be causes that only existences can be effects. In other words, the disappearance of motion *can* be an effect even if the disappearance of gasoline can *not* be considered a cause.

While the transition to a state of emptiness for the gas tank is causally relevant to explaining

---

[68] Ibid., 53

[69] Sterling P. Lamprecht, "Of a Curious Reluctance to Recognize Causal Efficacy," *Philosophical Review* 39, no. 4 (1930): 411.

[70] Ducasse, Types of Necessity, 135.

[71] Ibid.

[72] Lamprecht, "Curious Reluctance," 412.

[73] Ibid.

why the car stopped moving, the action of "disappearing" has no causal *efficacy* on Lamprecht's account because the "disappearing" is not a change in any *thing*. For an event to be a cause, the *event must be a change within a persistent entity*. If no such entity underlies the change, then the event may be referred to as a change in a situation or a change in a state-of-affairs, but not as an *action*. For an event to be causally efficacious, the event must be an action, *i.e.*, a change in the properties of a persistent entity. The disappearance of the gasoline was not a change in the gasoline, but the end of the existence of the gasoline then and there. Its disappearance was an event to be sure, but not a causally efficacious event. There is a non-causal relation between the disappearance and the stopping—a coincidence that is a sign of causality but not identical with it.

Lamprecht's analysis can be extended to other cases. He would deny, for example, that the disappearance of oxygen from the room caused the fire to go out, and that the disappearance of his son caused the father to become despondent. Saying that the disappearance of the gasoline cannot be the cause because the disappearance of the gasoline is not technically a persistent object misses the point. What *is* the persistent entity in this case?—The car. The event (loss of gasoline supply) is a change in a persistent entity (the car) that thus, on Lamprecht's own basis, has causal efficacy.

Let us consider the facts of the case. The capacity for the car to move is dependent on several factors, including:

- being in overall working mechanical condition
- presence of operator
- availability of gasoline supply to carburetor
- sufficient friction between tires and road surface

When a car accelerates from rest, its causal power is sufficient to overcome the resistance of the friction exerted upon its wheels by the road surface. Indeed, unless the car encountered adequate resistance from the road, it would not be able to exercise its capacity for movement at all—the tires would just spin as if the car were stuck on an ice patch. In other words, the resisting friction force of the road upon the car is a necessary condition for the exercise the car's power of locomotion.[74] Now, if the friction force of the road upon the car is a necessary condition for the exercise of the car's power—it cannot become *sufficient* for the stoppage of the car. It could not, therefore, properly be identified as *the* cause of the stopping. So it would seem that Ducasse wins this round against Lamprecht.

But does he?

If the presence of gasoline is a necessary condition for the exercise of the car's powers (which it has in virtue of it internal combustion engine and transmission system) just as the road friction is, it cannot become a sufficient condition for the car's stopping, and (given Ducasse's conception of causality) so the disappearance of the gasoline is not *the* cause of the car's stopping either.

E. H. Madden's answer is the correct one. He maintains, "running out of gasoline *and* the friction of tires on pavement, in turn, explain why the car stopped."[75] Since any exemplification of causality involves an entity acting in accordance with its intrinsic nature, *and* in accordance with its external conditions, it would be a mistake to try to argue that one or the other factor is entitled to metaphysical primacy, or that in statements of singular causal relations, there are *a priori* constraints

[74] Roads are an interesting example of entities with dual inverse capacities, as they have the capacity to enable and prevent the movement of vehicles upon them.

[75] Madden, "Hume and the Fiery Furnace," 65.

on the ontological category of existents denoted by the cause term in the statement.

§3. Events and Their Subjects

Hume himself was unclear about the ontology of the referents of singular causal claims. One of the unsettling features of Hume's arguments for the constant conjunction account of causality is his switching back and forth between employing the term "object" and the term "event" to specify the things related by causality. In the *Enquiry*, he uses the example of vibration causing sound to illustrate his conception of causality. First, he gives the general principle:[76]

> A cause [is] an object, followed by another, and where all the objects, similar to the first, are followed by objects similar to the second.

Next, he gives the example of how we interpret "this vibration caused this sound:"

> This vibration is followed by this sound, and that all similar vibrations have been followed by similar sounds.

Are vibrations *objects*? No. There seems to be no deeper explanation for this choice of synonyms than literary perversity. At any rate, standard exegetical practice among Hume scholars is to see "event" in most places where Hume says "object."[77]

While there is considerable debate among Humeans what events are, their criteria of identity, and so on, the standard assumption governing Humean approaches to the theory of causation is that the relata of the causal relation are events, and the terms involved in singular causal claims are names for events. Events, narrowly construed, are *changes* (with respect to properties) of things. More broad characterizations of events include "unchanges" as well, although if changes *and* unchanges are events this seems clearly too broad. A slightly more reasonable approach would be to say that event names are names for changes or unchanges at a time. Still, for an entity that is unchanging during the interval [t, t'], it is hard to see *what* makes an event corresponding to this interval non-trivially different from the event corresponding to the interval [t+ε, t'-ε], where $\varepsilon \ll |t-t'|$. Let us grant, then, that events are changes.

The causal realist believes that causality is a real process—a phenomenon of nature that characterizes how mind-independent material entities with particular natures act. This suggests very strongly that the causal realist must be, first and foremost, a metaphysical realist, by which I mean simply one who affirms the real existence of mind-independent objects. As I showed previously (in Chapter One), the standard sorts of considerations, such as the argument from perceptual relativity, offered in favour of rejecting Reidean common-sense realism, are groundless. In the absence of any compelling reason why we should deny the evidence of our senses, we are naturally in the position of realists for whom mind-independent entities can be perceived directly. The Humean theory requires an ontology of nothing much more than events.[78] The metaphysical realist countenances, in addition, entities with properties, changes in which *are* the events. The existence of events therefore *presupposes* the existence of both properties and entities. Just as there could be no dancing without a dancer, there could be no dancing without changes in the relative spatial orientation of the parts of the dancer's body.

---

[76] These are adapted from Hume, *Enquiry*, 51.

[77] There are, of course places where Hume really meant object in the sense of middle-sized physical object, but that is not our concern.

[78] Since the events are, on the Humean theory, subjective phenomena of consciousness, *mind* is also an implicit component of the basic Humean ontology. Hume's frequent references to "objects" refers to phenomenal events as objects of awareness, not to the extended material objects of common sense.

The term "event" is generally used to signify a change in state-of-affairs, but some philosophers (e.g., J. Kim) include both changes and "unchanges" in the category of events. Events may also fail to have subjects, such as the event of its becoming year 1996 C.E. "Action," unlike "event," has the unequivocal implication that it involves a subject, and that actions are changes in states of the subjects that undergo them. Usually, actions are regarded as that species of events in which the event is consciously intended.[79] If so, then plants do not act when they bend toward sunlight, acids do not act when they dissolve metals, nor can it be said that a pre-release version of a software application acts erratically! There is no purpose served by this usage, nor does it conform to the linguistic practices of scientists, engineers, *etc.*

One theory of events with recognizes the complexity of events (such as described in §3.2.2), and recognizes that changes are changes *in particular entities* has been advanced by Lawrence Lombard.[80] Since it takes the existence of objects as a given and analyses events in terms of changes in properties of objects, the account is consistent with a realist theory of causality.

Lombard identifies events as dynamic properties. The account of events distinguishes between simple and complex events, explicated as a distinction between "atomic" and "nonatomic" events. An atomic event is a specific change in a single entity during a specific period of time. Let us assume that an event e_1 causes another event e_2. Let $e_1 = (X,p,q,t,t')$ and $e_2 = (Y,r,s,t^*,t^{**})$, where X and Y are names of entities involved in the causal relation, while p, q, r and s are predicates signifying properties of the entities involved and t is a time parameter. Accordingly, X's changing from

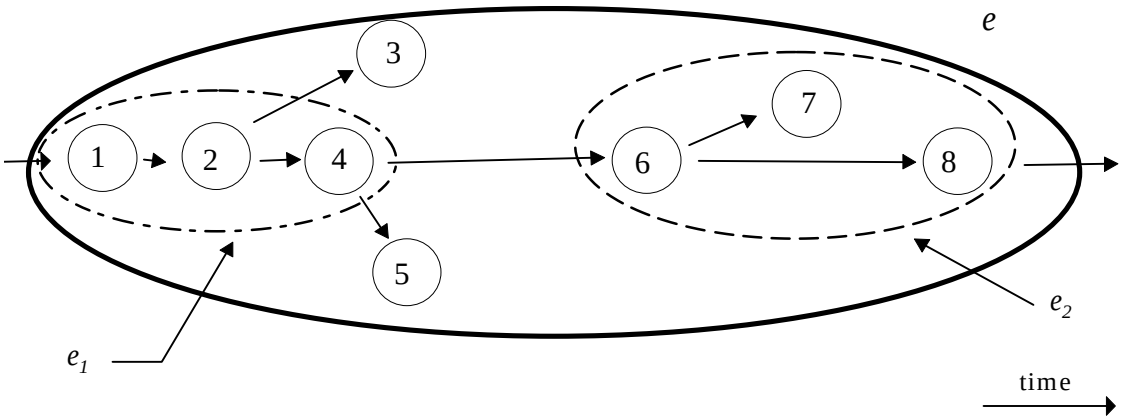

Figure 1: A Complex Event[81]

being p at t to being q at t' causes Y's changing from being r at t^* to being s at t^{**}.[82] X and Y are entities that have many properties, most of which remain constant throughout the changes that

[79] Johnathan Bennett, *Events and Their Names* (Indianapolis: Hackett Publishing Co., 1988), 188.

[80] Lawrence B. Lombard, *Events: A Metaphysical Study* (London: Routledge & Kegan Paul, 1986).

[81] In this diagram, arrows denote the transmission of causal influence, so event 1 causes event 2 is represented ①→②.

[82] $t<t^*$ and $t'<t^{**}$.

constitute the cause and effect events.[83]

To illustrate the theory, consider event *e* (Stanley's shooting Stella). It is, in Lombard's terms, a complex nonatomic event which has component atomic events. Some of these constituent events are: (1) Stanley's finger pulling the trigger, (2) the gunpowder exploding, (3) the propagation of a high-amplitude sound wave heard as a loud "bang," (4) the bullet leaving the chamber at a high speed in Stella's direction, (5) the gun recoiling, (6) the bullet entering Stella's flesh, (7) Stella's falling to the ground, (8) Stella's heart stopping. A complex event like this has its own internal structure which includes simultaneous constituent events having their own common causes, spatiotemporally contiguous cause-effect sequences making up causal processes, events that have no other effects, *etc.* (See Figure 1)

This complexity explains why we might ascribe various *properties* to events: the properties of a nonatomic event are dynamic properties of the objects involved in the complex event. Again, suppose we regard *e* as being constituted by two simpler events: Stanley's firing at Stella, (e_1) and Stella's being fatally wounded (e_2). We might describe e_1, as close-range, from the barrel of a Colt .45, or premeditated.[84] But suppose we say, "Stella was killed by a loud shot." It would be more precise (albeit more pedantic) to say, "one of the constituent atomic events of the shooting was the rapid oxidation of the gunpowder within the bullet's casing. This oxidation was the common cause of two events: namely, the propulsion of the attached bullet at high speeds towards its eventual victim, *and* the production of a high-amplitude sound wave in the surrounding air which was audible as a loud noise." Although the shot that killed its designated victim was loud, it did not have the effects *because* it was loud. A Colt .45 equipped with a silencer would have been just as deadly.[85]

Let *e* be a nonatomic event, constituted by a set of token atomic events e_1, e_2 … e_n such that $e = \cup\{e_i\}$. *e* has the property *F* iff there is set $\theta \subset e$ such that $F = \theta$. Informally, event *e* has the property of being *F* iff one or more of its constituent events was the dynamic property *F* of one of the objects involved in the event. In our example, to say that the shot had the property of being *loud* is just to ascribe a dynamic property "loudness" to the constituent atomic event (3), which was a dynamic property of the surrounding *air*. On the other hand, if the event is an atomic event, then the property *F* of an event c involving atomic object *X* is the dynamic property of the event, which is identical to *X's* having one static property at time *t* and another static property at time *t'*.[86] To summarize: an event is *F* if either (1) *F* is a dynamic property of a mereological component of a complex event, or (2) a (dynamic) property of an atomic object.

A cause is any change in an entity that produces a subsequent change in itself or another. Let $e_1 = (X,p,q,t,t')$ and $e_2 = (Y,r,s,t^*,t^{**})$, where *X* and *Y* are names of entities involved in the causal relation, while *p*, *q*, *r* and *s* are predicates signifying properties of the entities involved and *t* is a time parameter. Then "e_1 caused e_2" means "X's changing from being *p* at *t* to being *q* at *t'* causes *Y's* changing from being *r* at t^* to being *s* at t^{**}." Causality names the fact of its being the case that e_1 is

---

[83] If this were not the case, it would make no sense to say that *X* and *Y* continue to exist despite the changes occurring in each of them.

[84] Note that this simpler constituent event is *not* fatal, that is, none of its atomic constituent events is Stella's death.

[85] Ernest Sosa, "Mind-Body Interaction and Supervenient Causation," in *Causation and Causal Theories,* Midwest Studies in Philosophy, vol. IX, ed. Peter A. French, Theodore E. Uehling, Jr., and Howard K. Wettstein (Minneapolis: University of Minnesota Press, 1984), 278.

[86] Cynthia Macdonald, *Mind-Body Identity Theories* (New York: Routledge, 1989), 124.

causing e_2.

While it is possible to give a detailed account of events as Lombard does, and give the account a thoroughly realistic interpretation, the feature of the account that is striking is that non-atomic events have *internal causal structure.* What I refer to as a causal process can be understood in terms of this account of events. Causal processes can be regarded as the actualization of entities' powers to act and react in specific ways under specific circumstances—to do and to undergo motion and change through the operation of generative mechanisms. Clearly then, a realist theory of causation will require more metaphysics. However, as I hope to show, in the following chapter, the additional metaphysics does not introduce "spooks" such as "occult" powers, nor does it postulate in-principle unknowable "secret springs and principles" that must be regarded by the empiricist as spurious and arbitrary.

As I argue below, a realist theory of causation requires an ontology in which the basic components include: natures (including their compositional and structural aspects), powers (capacities and liabilities) which those natures carry, and conditions (including inhibitors and catalysts) necessary for the realization of capacities and liabilities. I regard causation as a process of power-actualization in an entity, and I view powers as determined by entities' intrinsic natures. In light of the conceptual dependency of causation upon powers, and of powers upon natures, these concepts demand explication in the reverse order. First, the concept of "nature" will be developed and arguments for including natures in our ontology presented. Next, the concept of a "power," especially its two variants—capacity and liability—will be explained and arguments for introducing capacities and liabilities into our ontology will be marshaled. Only after that groundwork is finished will we be in a position to articulate and defend a realist theory of causality.

CHAPTER FOUR: THE ONTOLOGY OF CAUSALITY

In Chapter 2, I discussed the general problem of the origin of our ideas of causality, and defended the position that it has its roots in the awareness of personal volitional efficacy—an awareness of the self as agent. Yet, a more general understanding of causality—one compatible with the accumulated knowledge and experimental practice of modern science—must account for the mode of causality involved in the actions of non-living entities as well as of persons and other living entities. In Chapter 3, I suggested that the most promising approach to causality involved the notion of entity-causation, a generalization of the agent-causation, in which causality is understood as a process of power-actualization rather than as a relation between events. In this chapter I seek to flesh out this suggestion by providing a case for the validity of the concepts of causal powers and natures.

This chapter documents two kinds of arguments that can be made for the validity of the concepts of powers and natures. The first, the "ontological arguments," move from very general metaphysical considerations to the validity of powers and natures, while the second, the "methodological arguments" suggest that the methods of causal inquiry in experimental science cannot be adequately explained without these concepts. For instance, on the first page of *Nature's Capacities and their Measurement,* Nancy Cartwright maintains that "we learn what the world is like by seeing what methods work to study it and what kinds of forecasts can predict it." She argues that "our typical methodologies and our typical applications, both in the natural sciences and in the social sciences, belong to a world governed by capacities, and indeed cannot be made sense of without them."[1] Harré and Madden's main philosophical goal for their book, *Causal Powers*, is implied on *its* very first page: "the construction of a conceptual system capable of accommodating the actual intellectual practices of science, and in which the known character of the world can be … systematically described."[2] As its title suggests, the notion of a causal power is a key one in their conceptual system. They are particularly clear in who their intellectual adversaries are; to them, there "can be no doubt that the Humean conception of Causality … must be wrong."[3] Cartwright is equally explicit that her "position is opposed to the tradition of Hume."[4] Seldom have philosophical battlelines been drawn so forthrightly and bluntly over questions of fundamentals.

The number and scope of the arguments I present shows beyond a reasonable doubt that insofar as science is concerned with causal questions, science makes little sense without the richer conceptual framework realism provides. In most branches of science, causal questions are focal. Scientists in various branches of inquiry are currently seeking to discover or confirm hypotheses concerning: the cause of the depletion of atmospheric ozone, the cause of the extinction of the dinosaurs, the causes of cancer, the causes of AIDS, the causes of anorexia, the causes of the business cycle, the causes of drought, the causes of clinical depression, the causes of psoriasis, the causes of tornadoes, and so on.

I first say what the concepts of powers and natures refer to (§4.1), and then present the ontological and methodological arguments on behalf of them (§4.2 and §4.3, respectively). Lastly,

---

[1] Nancy Cartwright, *Nature's Capacities and Their Measurement.* (Oxford: Clarendon Press, 1989), 2.

[2] Harré and Madden, *Causal Powers*, 1.

[3] Ibid.

[4] Cartwright, Nature's Capacities, 2.

(§4.4) I turn to the task of analyzing the logic of the arguments and assess to what extent they succeed at securing the additional ontology needed by the realist to present an account of causality as a process of power-actualization.

4.1 Realism, Natures and Powers

§1. Natures: Aristotelian and Modern

What residual conception of "nature" is necessary within the context of contemporary empirical science? We begin with the idea of a nature as identifying the *persistent, intrinsic structural*[5] *and compositional features of particulars* which play a causal role in determining the form and modes of alteration to which particulars of a given natural kind are liable, and whatever characteristic behaviors they are observed to engage in. For example, the nature of water is not its dispositional or physical properties, since they can vary, *e.g.* with temperature. The nature of water is just its molecular structure, since that is the only thing about water that remains invariant through its various changes of phase. The nature of *homo sapiens* and other biological kinds is their genotype.[6]

This conception of a nature preserves much of the sense of Aristotle's own theory of natures as presented in *Physics* II.1 and *Metaphysics* Δ.4, in a way that nullifies any serious basis for allegations that it countenances "occult" qualities; natures must be in principle empirically discoverable. Its sense is somewhat like that of Aristotle's "substance." We might say that a nature, like Aristotle's substance, consists of a compositional element (its material cause) and a structural or configurational element (its formal cause).

In her "Aristotelian Natures and the Modern Experimental Method," Nancy Cartwright presents arguments on behalf of the reality of "natures." There are hints that she is not completely consistent in her conception of what these "natures" are, however. At one point, she maintains that "Newton's conclusions … throughout his work on optics, are about the inner constitution of light. I claim that this study of the inner constitution is a study of an Aristotelian-style nature."[7] Yet at another place, she maintains that "natures are something like powers."[8] I contend that Cartwright is correct to refer to an inner constitution as a nature ("Aristotelian-style"), but saying that a nature is "something like" a power is as misleading as it is vague. If natures are the inner constitutions of things, whereas powers are potentialities that are had by things *in virtue of* their natures, then powers are *not like* natures—powers are determined by natures. In those contexts where Cartwright makes this lapse, what she is describing is what I refer to as "liabilities." A liability is a disposition to be *affected by* the action of some agent. It is true that an entity with a certain liability has, "in its nature," the power to respond to the action of another, but a liability is not "something like" a nature. Natures, as Cartwright agrees, are not irreducible subsisting essences but intrinsic structural features of entities.[9] They are, as Aristotle maintained, "causes and principles of change" that give rise to characteristic modes of behavior, and those modes of behavior will be *necessarily connected to those structures*, so long as the structures are what they are: "when we associate a particular princi-

---

[5] For living things, an important type of structural feature—that which realizes *functions*—needs to be identified for teleological explanations.

[6] See Harré and Madden, *Causal Powers*, 103-109.

[7] Cartwright, "Aristotelian Natures," 69.

[8] Ibid., 48.

[9] Ibid., 47.

ple of change with a given structure or characteristic, we expect that association to be permanent, to last so long as the structure is what it is."[10]

The element present in modern science that is missing in Aristotelian science is the element of intervention. In order to discover natures, we will have to dig under the surface of phenomena, since the invariants we are looking for in the physical sciences today do not manifest themselves in observable regularities. If we take "actions" to be the objects of explanation, we can divide the causal determinants of actions into two categories: intrinsic and extrinsic determinants. The intrinsic determinants of actions correspond roughly to Aristotle's formal, material and final causes, while the extrinsic determinants include efficient causes. The "nature" of a thing is constituted by its intrinsic, non-relational properties which are associated primarily with its structure and composition. The capacities and liabilities—*i.e.,* the "powers"—of a thing are the relational properties carried by an entity in virtue of its specific nature.

Particulars have powers, which they possess in virtue of their intrinsic natures. Powers represent kinds of potentialities for action and for alteration that are actualized when the conditions for such are in place. Powers come in two types—capacities and liabilities (dispositions). "Capacity" refers to an active power—the potentiality for a particular to produce a change or alteration in some other entity, while a liability is a passive power—the potentiality for a particular to undergo a change or alteration under the influence of some other entity. "Causality" names the *process* of power-actualization, a process in which an entity, given a specific set of extrinsic conditions, produces and / or incurs some specific change. To say that a particular has certain powers "in virtue of" its intrinsic nature is to maintain that there is a connection between the natures of things and their powers such that the nature of a thing *realizes* and *determines* the powers that it has. I contend that scientific explanation needs natures, and that the ontology of natures needs to be articulated from an Aristotelian perspective.

To show the historical methodological continuity between Aristotelian and modern science, Nancy Cartwright[11] gives qualified endorsement to Francis Bacon's (1561-1621) conception of natures: "Bacon still employs the Aristotelian idea of natures and essences, but for him these are not hidden. Bacon looks for the explanatory essences, but looks for them among qualities that are observable."[12] She explains how, by eliminative induction, he correctly identified the nature of heat as the motion of minute particles in bodies, and established the basis for the kinetic molecular theory of heat. The claim is then made that explanation presupposes identifying the natures of things even in modern scientific contexts:

> … modern natures are like Bacon's and unlike those of the Scholastics, in that they are attributed to observable structures and qualities. Generally they differ from Bacon's in that they do not lie on the surface and are not to be observed with the naked eye. Rather, we often need very subtle and elaborate experiments in order to see them. Modern science insists that we found explanation on experimentally identifiable and

[10] Ibid.

[11] Ibid., 46.

[12] Ibid., 46. The most generally accepted ancient explanations of heat were compositional: heat was associated primarily with the presence of the element fire. Heat as the motion of atoms in the void must have seemed extremely counter-intuitive to those for whom compositional notions of physical explanation were paradigmatic of good explanations.

> verifiable structures and qualities. But, I maintain, what we learn about these structures and qualities is what it is in their natures to do.[13]

There is nothing either "occult" or anachronistic about natures when they are conceived as I have suggested they should be. The same applies to the "powers" that things possess in virtue of their natures.

A "power" of a thing or substance is any potential for action or change which it is capable of doing, or to which it is liable, but whose actualization is dependent on the presence and/or absence of specific conditions in its environment. The metaphysical perspective of causal realism regards active entities as the primary particulars of the world or as Rom Harré and Edward Madden dub them, "powerful particulars."[14] A powerful particular is an entity with a determinate material composition and structure—*i.e.*, a specific nature—which acts and interacts with other entities in ways strictly consistent with what its nature makes possible. According to causal realism, to say that the ways in which an entity can act and interact with others is determined by its nature means that the nature of the entity determines and realizes the powers of that (token) entity.

> The proper analysis of the ascription of a power to a thing or material … is this: 'X has the power to A' means 'X will do A, in the appropriate conditions, *in virtue of its intrinsic nature*.'[15]

The distinction between a power and a nature, in simple terms, is this: a nature refers to what something essentially *is*, a power refers to what something essentially *does*—*i.e.*, does "by nature." As indicated, these invariants fall into just a few basic categories: biological genotype, chemical element or substance, or elementary particle. The "Lockean primary qualities" of things are still associated with the basic properties of the "insensible" (not detectable by the unaided senses) parts of things, but now the number, motion, arrangement, *etc.* of these insensible parts can be characterized much more precisely as valence, quantum spin, base-pair sequence, and so on. If the nature of a thing determines the range of the supervenient powers that things of that nature have, then that nature poses an ontic constraint on what types of actions are possible for entities which share that nature.

§2. Powers: Capacities and Liabilities

To say that "power are real," is to say that it is legitimate to subsume, under a single concept—*i.e.*, the concept "power"—both the occurrent action of a particular when it is acting, as well as the latent potential for just that kind of action which continues to exist, unactualized, when the extrinsic conditions necessary for such action to occur are unsatisfied, or there are countervailing factors inhibiting the action. Let us now survey the extant approaches to validating the concepts power, capacity and liability.

In the Lockean tradition, to ascribe a power to a material object is

> to assert what it can or cannot do in virtue of its intrinsic nature in relation to specifiable extrinsic circumstances, leaving open a complete characterization of the object's constitution in virtue of which it is held to be endowed with powers.[16]

---

[13] Ibid.

[14] Madden's acknowledgement of the influence of Averroes on his own conception of causation is noted in Kogan, *Averroes*, 284 n. 32.

[15] Harré and Madden, *Causal Powers*, 86. This version is slightly simplified, to bracket the modifications necessary to integrate human active powers into the theory.

[16] P. M. Heimann and J. E. McGuire, "Newtonian Forces and Lockean Powers: Concepts of Matter in

Harré and Madden express similar sentiments:

> To ascribe a power to a thing or material is to say something specific about what it *will* or *can do,* but it is not to assert any specific hypothesis about the nature of that thing. To ascribe a power to a thing asserts only that it can do what it does in virtue of its nature, whatever that is. It leaves open the question of the exact specification of the nature or constitution in virtue of which the thing, person or material has the power.[17]

Moreover, to ascribe a power is to say

> what the material does or can do is to be understood as brought about … in some measure by the nature or constitution of that thing or material, *i.e.,* by intrinsic conditions.[18]

Further, the authors identify the possession of power with the possession of agency, and distinguish a different sort of power that indicates passivity:

> It is through the distinctions between the degree to which intrinsic nature and extrinsic circumstances enter into an explanation of the behavior referred to in the power ascription that we make the distinction between agency and passivity. We shall generally use the term "liability" for a "passive power"….[19]

Based on the Lockean distinction adopted by Harré and Madden, I will generally refer to active and passive powers as "capacities" and "liabilities" respectively.

When an entity possesses the power to initiate a set of consequences or to cause an effect, that establishes the agency of the entity with respect to the specific powers it has, whether it is a sentient organism or an inanimate object. We say that the entity has the "capacity" to produce some effect. Examples of capacities are: the capacity of an acid to dissolve metal, of a man to create a sculpture, of a virus to induce lethargy, of caffeine to stimulate alertness, of an aspirin tablet to relieve a headache, of a field to accelerate a particle, of a brake to decelerate a motor vehicle, of a well-built building to resist hurricane damage. As Locke maintained, capacities are relational (secondary) rather than intrinsic (primary) properties, of entities.

When an entity possess the power to be altered or changed by the activity of some *other* entity, that indicates its passivity with respect to the specific action in question: it cannot resist or counter the activity that produces the change. We say in such cases that the entity has a "liability" to undergo the change. Entities that possess these liabilities are generally on the "receiving end" of the causal transactions in which they are involved—they are the objects that sustain damage, suffer alteration, get bumped aside, *etc.*[20] Examples of liabilities are: the liability of a body to be accelerated by a force, of a fragile object to break when dropped, of paper to burn or of dynamite to explode when ignited, of a plant to die when deprived of adequate light.

Powers are special kinds of relational properties because they are not always actualized in the

---

Eighteenth-Century Thought," in *Historical Studies in the Physical Sciences,* ed. Russell McCormmach (Philadelphia: University of Pennsylvania Press., 1971), 235.

[17] Harré and Madden, *Causal Powers,* 87.

[18] Ibid.

[19] Ibid., 89.

[20] The other terms, such as "tendency," "disposition," and "propensity," are surely also plausible candidates for inclusion in the lexicon of power-talk. They also surely all have connotations derived from their presence within other accounts of causality. To avoid importing theoretical baggage from incompatible approaches that might undercut the effectiveness of my own arguments, my positive account of causation will ignore them for the most part.

entities that have them;[21] they continue to exist as latent potentialities for action in entities when they are neither fully nor partially realized. Powers are sensitive to the presence of specific conditions for their actualization, and the absence of inhibitors of their action. For example, any object is liable to accelerate if it is subject to an imposed force, *given* the absence of countervailing forces acting to inhibit the motion it would otherwise experience. A fragile object has the liability to break when dropped, provided it is not dropped from a low height, nor that its landing is softened by an impact-absorbing surface such as a carpet.

Since the actions of entities manifest a kind of causal interdependence, the actualization of a power, whether a capacity or liability, will require that the "causal context" of the entity be "right"—that is, that the necessary conditions for the actualization of the power are present, and that inhibitors are absent. A "causal context" of an entity A, is the totality of all those entities within a given region of space-time S, centered on A, that are causally related to A. It is the sum total of all the existents conditioning the actions or interactions of any specific entity. The universe is not a scattered assortment of ontologically independent objects each of which stands apart from the rest. Nor, on the other hand, does the universe exhibit the sort of global systematicity that might warrant treating the universe as a integrated whole, or a single entity. Such talk would be poetic at best. Causal interdependency is a pervasive feature of the natural world, but it is *local* interdependency, from which emerges local systematicity.

For the Humean empiricist, talk of capacities is unnecessary metaphysics, because we can replace any sentence with the word "capacity" by another sentence with the word "law" or "regularity" without any loss of explanatory power or of meaning. As the following arguments show, that is exactly what you *do* lose. If we are interested in providing the best conceptual framework for the analysis of experimental reasoning, the inclusion of natures and capacities will prove to be indispensable. The realist's ontology provides a way to connect the explanation of observable phenomena to the structure of the entities investigated by science, and provides a richer framework for the analysis of the logic of science than that furnished by the Humean empiricist ontology of events and phenomenal qualities.

Human beings are complicated entities, whose biological nature endows them with a wide variety of natural needs—love, money, art, philosophy, and self-esteem—to name but a few. The satisfaction of all of these needs requires that the individual take specific actions to procure the means by which they may be satisfied. But individuals cannot be at two places at the same time. They must, at each moment, choose to act in one way rather than in any of the others possible, and somehow turn a diverse range of actions into a pattern of organized behavior consciously designed to be efficacious in value-acquisition. Our nature gives us the capacity to do many things, but not all at once. Over the course of a life-time, we can only do a fraction of the things of which we are capable. Even insentient, non-volitional entities like the simplest of molecules will only have the opportunity to interact with the matter in its particular corner of the universe, and so many of its powers will go unrevealed. Whereas the *nature* of any given thing is actualized at each moment of its duration in existence, the *powers* which that nature gives rise to are typically not.

[21] For Cartwright, however, capacities are not relational in the sense I mean.

4.2 Metaphysical Arguments for the Realist Ontology

§1. The Explanatory Inadequacy of Reductive Materialism

The existence of what Locke and Reid called "active powers," by which our own bodies are called into motion by internal thoughts, motives and passions, is almost as indubitable as the existence of our own minds. The existence of the power of agency is obvious, perhaps tediously so, in cases of spontaneity, like the one that Harré and Madden offer as a paradigmatic example:

> On a fine sunny afternoon of only moderate heat, and with no breeze to speak of, a man dozes in a deck-chair in a garden. There are no flies, mosquitoes nor wasps, nor shout from neighbours' children. Suddenly the man jumps up, walks smartly to the shed, takes out the lawn mower and begins to mow the lawn. Nothing extrinsic to him has changed. His subsequent actions are the products of his decision, the sources of which are to be found among states intrinsic to him.[22]

There is no extrinsic efficient cause of the man's behavior in this case. From the "outside," there is no efficient causation happening at all.[23] By contrast, insentient nature seems to be entirely passive.

A billiard ball, sitting on a billiard table will not move until bumped by another one, or until a cue strikes it—it has no power of locomotion. Efficient causation is sufficient to explain its motion. When the ball is struck, predictions can be made to a striking degree of accuracy about the subsequent motion based on simple principles of mathematical physics. There is no concept of active power that is needed in order to conceptualize the sequence of events involved in billiard interactions—the law of conservation of momentum, at least in principle, can subsume all the billiard ball activity heretofore observed, and by extension, to all the mechanical activity of impenetrable, inelastic, solids on flat, frictionless surfaces. In this case, ascribing an active power to a billiard ball would be both unnecessary as a methodological posit and arbitrary as a metaphysical one.

Could such an analysis be extended to cover more complex phenomena, such as the mental events of human beings? Some philosophers have responded in the affirmative, and have denied the reality of the active powers, maintaining, on the contrary, that human beings are just complex deterministic objects, passively responding to various mechanical forces acting from within and from without.[24] As I argued in Chapter 2, cognition is a capacity that requires volitional actuation; voluntary motor behavior is a capacity that requires conscious direction, so the active powers of persons have their ultimate source in volition.

The appeal of physicalistic materialism in the philosophy of mind is based on four interrelated assumptions: (1) the entire world (including the brain) is a causally closed system. Since (2) the causal laws governing the behavior of the world involve only physical event-type predicates, therefore (3) events described in physical terms are the only legitimate objects of explanation. Lastly, (4) efficient causation is the only mode of causation.

The causal closure of the physical world[25] can itself be regarded as a kind of conservation

---

[22] Harré and Madden, *Causal Powers*, 83.

[23] Of course, there is internal efficient causation occurring, but only at the level of **CN**S electrical activity, muscle contraction, and limb movement. No efficient causal explanation would explain the co-ordinated activity of the entire person, nor its ostensibly goal-directed character.

[24] Thomas Hobbes was one of the first materialists, followed later by thinkers such as Baron Paul d'Holbach and Karl Marx. The denial of the existence of a "soul" is a clear implication of materialism and thus materialism has historically been associated with atheism.

[25] Laird Addis, "Parallelism, Interactionism and Causation," in *Causation and Causal Theories*, Midwest Studies in Philosophy, vol. IX (Minneapolis: University of Minnesota Press, 1984), 329-44.

principle—no energy can enter the system from outside it, nor "leak out." This bars any irreducibly mental event from being a cause of any physical event, given that such an occurrence would imply transfer of energy to the *ex hypothesi* closed physical realm from a separate mental realm. Any mental event, such as a spontaneous desire to mow one's lawn, can be characterized as a mental event *or* a physical event (albeit a complex one). Indeed, the mental event *must* be capable of being so characterized, otherwise it would be completely inexplicable, given materialistic assumptions. The rub is that conscious mental activity is "nothing but" the activity of complex physical processes in the brain and central nervous system; mental and physical activities of the mind/brain are at least "token"-identical, if not locally "type"-identical.

Physicalistic identity theories also draw inspiration from the idea that the activities of both mereologically complex inhabitants of the physical world and their simple components are governed by the same mode of causality—and by the same *kinds* of causal laws. Efficient causality is the mode of causality involved in event-event causal relations, and the causal laws are those which express laws of invariable physical-event-type succession relations. Since the physical properties in virtue of which objects interact with one another the way they do in the system are the only causally relevant properties, only physical-event type predicates appear in causal laws.

The identity thesis, combined with the thesis that all causal processes occurring at the physical level in the brain and central nervous system are governed by causal laws tends to support the contention that consciousness itself is "nothing but" the activity of complex physical processes in the brain and central nervous system.

The fallaciousness of the argument should be readily apparent. Even if the premises were both true, which I am prepared to grant, the conclusion might still be false. In general we cannot expect composites to behave in the same way that their "ultimate" physical constituents do, just as we cannot in general infer that any given property of such constituents will also be true of the composite. The fact that consciousness is subject to volitional control, and that the brain and central nervous system is not, suggests that at some levels of organization, novel properties emerge, and entities have causal powers in virtue of having certain emergent properties.

Even if this admittedly hasty sketch of the motivation for taking causal powers of human beings as both self-evident, and *not* easily reasoned away by physicalism, it remains to be seen whether there is any good reason for accepting an ontology of powers into the world of the simpler entities of botany, chemistry, solid state physics, *etc.*

§2. Object Persistence

Harré and Madden invoke Locke's distinction between real and nominal essence, and based on this distinction, they affirm that one ought to

> identify the nature of a kind, material, or individual with its real essence; that is, what according to the most advanced scientific viewpoint of the day, is taken to be causally responsible for the properties and powers it manifests, and in particular for those properties which satisfy its identity criteria both qualitative and numerical.[26]

Harré and Madden's characterization of natures anticipates Cartwright's neo-Aristotelian conception, whereby natures "do not lie on the surface and are not to be observed with the naked eye. Rather, we often need very subtle and elaborate experiments in order to see them."[27] Likewise,

[26] Ibid., 101-2.

[27] Cartwright, "Aristotelian Natures," 46.

Harré and Madden affirm that they "will not be revealed to the unaided inspection of the senses. Things are physically and chemically more complex than their surface appearances and their inner properties are in general different from their manifest properties."[28] The concept of a nature presupposes a fine-structure conception of physical substances whose causally basic properties can be detected by sophisticated experimenters.

Is the distinction between the persistent, causally basic characteristics of entities and the alterable, causally derivative characteristics of entities a valid one? The distinction for which the concept of a nature is most obviously needed is the distinction between persistence and change within a particular. The very idea of there being a persistent existent which could be the bearer of change is presupposed by the concept of change. The distinction seems unintelligible without it. Consider what an action is: it is a particular kind of event—one involving an agent cause. When a particular acts, it undergoes some alteration or determinate change with respect to one or more of its attributes (or "characteristics") over a specific period of time.

> The continuity in existence of a member of a kind, or a sample of a material substance, is referred to its continuing to have the properties that constitute the real essence of the kind or material.[29]

To say that *an* entity undergoes a change presupposes that the entity is a continuant—that the entity survives the change with respect to one or more of its attributes. In other words, the entity after the change is token identical to the entity before the change, despite not being identical in every conceivable respect.[30]

The causal realists' metaphysics of powerful particulars presupposes that entities, to varying degrees and at various rates, undergo change. Particulars are inherently dynamic; change is a basic aspect of existence. Since change is a phenomenon of motion, it follows that the existence of motion is a brute fact of reality, not itself demanding explanation. But the characteristic kinds of change that things undergo and the dispositions they have to be affected by background conditions and other objects *do* demand explanation. The fact that in the same forest some kinds of trees will lose their leaves in the winter and others will not is hardly something that a botanist might reasonably take to be a brute fact not in need of explanation. A scientific attitude toward nature is one that seeks causal explanations for characteristic behaviors of physical and biological entities. The fact that entities exhibit certain *characteristic* behaviors and dispositions suggests that such behaviors and dispositions are causally related to those persistent aspects of the particular in question—the nature of the particular.

§3. The Analysis of Disposition Terms

In order to further contrast my approach with one logical empiricist one, and show how the introduction of the metaphysical concept of "nature" is warranted, consider the empiricist analysis

---

[28] Harré and Madden, *Causal Powers*, 102.

[29] Ibid., 102.

[30] According to Leibniz's Law, two particulars are identical if they have all the same properties:

$x=y \Leftrightarrow (F)(Fx \rightarrow Fy)$.

This condition clearly does not capture the sense of identity relevant to the analysis of action; a change with respect to one property of the entity would mean the end of that entity's existence and its replacement by some new entity that is the same in every way save one.

of the meaning of disposition terms.[31]

Disposition terms are those that name physical properties which are not instantiated at every moment of the existence of the object that has them. Disposition terms include "soluble," "malleable," "electric conductor," *etc.* According to the analytic tradition, the definition of a term should specify the necessary and sufficient conditions for its employment. Suppose we wanted to define the disposition term "fragile" in the following way: an object *x* is fragile if and only if at any time *t* the object is sharply struck, then *x* breaks at that time. The definition could be symbolized by

$$Fx \equiv (t)(Sxt \supset Bxt)$$

where '*F*' stands for 'fragile,' '*S*' for 'struck,' and 'B' for 'breaks,' and *t* is the time parameter. The problem with a full definition in terms of necessary and sufficient conditions emerges from the truth-functional semantics of the conditional. If some non-fragile object (like a raindrop) is never struck during its existence, then '*Sat*' is false and thus '$Sat \supset Bat$' is true for all values of *t*. Consequently '*Fa*' is true though *a* is not fragile. Carnap's solution was to introduce 'reduction sentences,' which differed from analytic definitions in that they would only give partial interpretations of the disposition terms. The term 'fragile' would be 'introduced' by the following reduction sentence:

$$(x)(t)\ [Sxt \supset (Fx \equiv Bxt)]$$

If *x* is struck, then it is fragile only if it breaks, and nonfragile if it does not. However, if *x* is never struck, the truth value of *Fx* is indeterminate rather than true—this eliminates the earlier difficulty. In other words, if the object is never struck, we cannot assign a truth-value to the assertion '*x* is fragile.' If the object is never struck sharply and therefore never has the opportunity to satisfy the condition of its fragility, (*i.e.*, its breaking), then the term 'fragility' itself is rendered devoid of empirical content.

Hempel acknowledges that the best that can be done for disposition terms is this "partial" or "conditional" definition. It is nonetheless presented as a "satisfactory interpretation of the experiential import of a large class of disposition terms."[32] If this reduction sentence is the best that can be done as an analysis of the meaning of disposition terms like 'fragility,' the analytic empiricist approach is surely vacuous. To see why, suppose we have two virtually identical glass vases (same physical dimensions, made from the same slurry of oxide, same weight, *etc.*) sitting on a table. Call them vase *a* and vase *b*. Suppose vase *a* is struck sharply and thereupon broke, while vase *b* was left intact. Assuming the truth of (R), we would infer two things. First, we would infer *Fa*—vase *a* would be known to be fragile (at the moment of breaking). Second, we would have to infer that vase *b*'s fragility is unknown, since the truth value of '*Fb*' is indeterminate if ~*Sat*.

To most of us, this second inference seems clearly invalid; given the information known, we believe that we are rationally justified in accepting the conclusion that vase *b* is fragile. The conditional analysis of disposition terms is incapable of showing us how to express an inference that to most of us is plainly sound. The empiricist might object that the inference to the fragility of *b* is not logically warranted at all—that it is just a manifestation of our habituation to expect similar consequences from similar antecedents. If this were true, then we should experience the act of making such an inference as being more intuitively compelling the more repetitions of the appropriate habituating sequences one observes. But we do not. On the contrary, it makes *no difference* whether one has seen two or two hundred vases break, what is relevant is the material type-identity of the

[31] This summary follows Carl G. Hempel, "Empiricist Criteria of Cognitive Significance: Problems and Changes," in *Aspects of Scientific Explanation* (New York: The Free Press, 1965), 109-110.

[32] Ibid.

two vases *a* and *b*. Why do we feel that we *know* that the second vase is also fragile?—Because we accept the validity of this argument:

(P1) Vase *a* is fragile. (given)

(P2) Vase *b* is identical to vase *a* in all essential respects. (given)

(P) Therefore, vase *b* is fragile.

This is an expression of a more general pattern of empirical reasoning that we take for granted:

(E1) Entity *a* under conditions *c* incurs change *x*.

(E2) Entity *b* is identical to entity *a* in all essential respects.

(E) Therefore, entity *b* under conditions *c* incurs change *x*.

As stated, this argument is not logically valid; it is enthymematic. What, then, are the suppressed premises? I submit that they are two *metaphysical* assumptions: (i) that things which are identical in all essential respects have the same nature, and (ii) the same powers are associated with the same natures. These assumptions are implicit auxiliary hypotheses whose explication is required if the conclusion (E) is to be adequately justified. When these premises are added to the argument, it becomes as rigorous as can reasonably be expected given the context:

(E1) Entity *a* under conditions *c* incurs change *x*. (given)

(E2) Entity *b* is identical to entity *a* in all essential respects. (given)

(E3) Any two entities that are identical in all essential respects have the same nature. (assumption I)

(E4) Any two entities that have the same nature have the same liabilities [powers in general]. (assumption II)

(E5) Any two entities that have the same liabilities incur the same change under the same conditions. (from def'n of 'liability')

(E) Therefore, entity *b* under conditions *c* incurs change *x*.

The first assumed premise (E3) involves a commitment is to the existence of natures. The nature of an entity is the set of properties it possesses. The totality of an entity's essential properties *constitutes* its nature. The distinction between essential and accidental properties was originally intended to differentiate the stable, invariant aspects of an entity's constitution from the variable, temporary features the entity may lose or acquire during its existence. Two entities that have the same essential properties are entities that have the same nature. If two things are made from the same material by the same process, as two vases blown by an expert glass blower, we assume that the aspects of their constitution that are causally relevant to their actions are identical.

The second missing premise (E4) involves a commitment to the existence of a connection between aspects of a thing's nature and its various liabilities (a parallel assumption would hold for capacities.) We expect that both vases will break because they are both fragile, and that they are fragile even when they are not being broken for the *same* reason: we suppose that their common liability to break is a direct consequence of the impact-sensitivity of their constituent crystal structures. The only available justification for our belief that the second vase is also fragile is grounded on a commitment to the identity *in rerum natura* of the liabilities of two things that have

the same nature. Two things that are of the same nature will act identically under the same conditions.

If (E3) - (E5) are true, then the idea that the meaning of disposition or liability terms can be specified conditionally is false. It follows that Harré and Madden are correct to make the following comment:

> Things and material *have* powers even when they are not exercising them, and this is a current fact about them, a way in which they are currently differentiated from other things and materials which lack these powers…. The difference between something which has the power to behave in a certain way and something which does not have that power is not a difference between what they *will do*, since it is contingently the case that their powers are, in fact, ever manifested, but it is a difference in what they themselves now are. It is a difference in intrinsic nature.[33]

4.3 Methodological Arguments for the Realist Ontology

In her influential *Nature's Capacities and their Measurement,* Nancy Cartwright offered an extended case for the introduction of a realistic conception of causality, according to which "capacities and causings are real things in nature."[34] According to Margaret Morrison, "Cartwright presents a compelling case for the role of capacities in both natural and social science. Her examples strengthen the philosophical points in ways that take us well beyond traditional metaphysical arguments for capacities or dispositions."[35]

Cartwright's overall project is to show how empiricism in the philosophy of science can and *should* assimilate a realistic metaphysics of causality, provided that empiricism is conceived of as being defined by concern with questions of testability and measurability, not meaning and axiomatization. While I am sympathetic to the overall project, my interest in her work for the purposes of this chapter is quite delimited. I want to identify the best arguments she offers for the validity and indispensability of capacities, clarify the logical strength and structure of those arguments, and sharpen the contrast between the Humean paradigm and the realist alternative Cartwright pictures. To achieve these ends, I have not needed to present the technical details of the arguments, which can be got by reading her work. I have, however, needed to modify the language slightly at several points in order to achieve a uniformity of expression with other chapters of the present work. The reader should not assume, therefore, that my usage of the terms "capacities" and "tendencies" will correspond exactly to Cartwright's. By assimilating Cartwright's arguments for a realist view of capacities into the larger project of expounding a systematic version of causal realism, my argument benefits from the strengths of her own, while her arguments benefit, in turn, by being cast as an important part of the larger project to supplant the empiricist theory of causality.

§1. The Argument from Complexity

The philosophical backdrop against which the causal realist's project should be viewed is provided by the metaphysics of Aristotle and John Locke. Yet, the realist's debt to history could hardly be paid without giving credit to John Stuart Mill, at least insofar as the arguments he offered for the validity of "tendencies" anticipates the modern case for capacities. Cartwright, in particular,

---

[33] Harré and Madden, *Causal Powers*, 86.

[34] Cartwright, *Nature's Capacities*, 170.

[35] Margaret Morrison, "Capacities, Tendencies and the Problem of Singular Causes," *Philosophy and Phenomenological Research* LV, no. 1 (1995):163.

thinks that Mill is motivated by the right kind of concerns.

The context for Mill's introduction of tendencies is the analysis of the logic of the "moral sciences," including what we now call the "social sciences." The problem is that the nature of social phenomena seems to preclude induction to the kinds of generalized causal laws that we assume are necessary for explanation and prediction. Social phenomena are the results of the confluence of complex and constantly shifting sets of factors, no particular arrangement of which is stable enough over time for us to extract the empirical evidence of their consequences necessary to generalize from.

Here is a pedestrian illustration of what happens when we try. Suppose we have the following (true) generalization:

> (**G**) Four male friends play golf *every* time the following conditions are satisfied; that it is a Sunday in June, July or August, that the probability of precipitation is predicted to be less than 50%.

Now, suppose on a sunny Sunday afternoon in July, we find Tom painting his deck, Dick working on the truck, Harry surfing the Internet and Homer teaching his daughter how to ride a bicycle. What happened? Homer's wife Marge insisted that he was spending too much time with his golfing companions and not enough time with the kids, so Homer called to cancel the tee-off time. Of course, we could respond that **G** is only true *ceteris paribus.* But suppose this breakdown in the golfers' arrangements is the norm rather than the exception—that the *ceteris paribus* assumption is not satisfied most of the time. What sense does it make to say that a generalization is true, all other things being equal, *if* other things seldom *are* equal? The generalization is pretty trivial, and not very useful, to say the least.

Cartwright agrees with Mill that the laws of mechanics are similar, despite the fundamental difference in the nature of the entities involved. Consider another (true) generalization:

> (**G***) Any object dropped from above the surface of the Earth accelerates along a trajectory towards the center of the Earth at a rate of 9.8 ms^{-2}.

This too is a *ceteris paribus* law. But even more remarkably, this condition almost *never* holds, for it is literally true only if the object is released inside a vacuum, so that there is no air resistance. In a similar context, Mill suggests:

> We might, indeed, guard our expression … by saying that the body moves in [the prescribed] manner unless prevented, or except in so far as prevented, by some counteracting cause, but the body does not only move in that manner unless counteracted, it *tends* to move in that manner even when counteracted.[36]

For Mill, causal laws are not laws about what things actually *do* but what they *tend* to do, because what things actually do is to at least some extent situation-sensitive in a way that makes generalization strictly speaking, impossible. What we *can* confirm is that things have a *tendency* to act in certain ways. Mill's causal laws are thus laws about tendencies, which are to be distinguished from causal generalizations that state partial or counterfactual regularities in the normal course of events in the world. The difference between Hume and Mill here is stark: Mill thinks the first kind of laws are the true causal laws, Hume thinks the second kind are.

Tendencies provided a way to preserve the universality of the causal laws describing some characteristic action of entities of some type, despite the presence of factors that might inhibit any

---

[36] John Stuart Mill, *A System of Logic, Ratiocinative and Inductive,* 8th ed. (New York: Harper and Brothers, 1874), quoted in Cartwright, *Nature's Capacities,* 177.

token from doing what it otherwise would. If the syntax of a causal law was a capacity ascription to an entity-type rather than a *ceteris paribus*-hedged counterfactual conditional, then the universality of the laws implies its *exceptionlessness*. Mill draws a profound implication from this:

> What is thought to be an exception to a principle is always some other and distinct principle cutting into the former, some other force which impinges against the first force and deflects it from its direction. There are not a *law* and an *exception* to that law—the law acting in ninety-nine cases, and the exception in one. There are two laws, each possibly acting in the whole hundred cases and bringing about a common effect by their conjunct operation.[37]

Mill's "tendencies" seems very close in meaning to what I have been referring to as "powers," while Cartwright and I use the term "capacity" to pick out just those powers that bring about something, or produce some effect—those that signify the causal agency of the particulars whose powers they are.[38] Laws about tendencies are, for Mill, laws about stable or fixed capacities, not laws about regularities.

In short, the argument to capacities from complexity is this: phenomena in the world are complex conjunctions of causal factors. This complexity makes generalizations about the actual behavior of individuals impossible to sustain without the addition of *ceteris paribus* clauses, most of which are seldom if ever satisfied. Simple induction by observation never gets at any laws that are simultaneously generalizable, true and interesting. But science needs universal, exceptionless generalizations for prediction and explanation. Therefore, science needs laws about tendencies.

Cartwright's argument for capacities derives from a well-founded rejection of the Humean assumption that nature wears her regularities on her sleeve. There are, of course, well-known inadequacies of this view, but it has nonetheless been regarded as a starting point for philosophers of science in the Humean-empiricist tradition. If Cartwright is right, then the view should be rejected entirely, used not even as a starting-point. As I suggested earlier, causal realism should take the Aristotelian view that entities are mereologically complex. It should also adopt the Mill-Cartwright view that nature is complex. Nature will only reveal her laws to us under specially contrived conditions which occur typically in controlled experiments. Nature can only be made simple by human intervention—her laws are discovered only after regularities are *created*. This view is, I think, one that the causal realist ought to adopt.

§2. Justifying Induction: Newton's Experimentum Crusis

Cartwright interprets Newton's *experimentum crusis* as reported in his 1671 letter to the Royal Society (and in his later optical writings) as one that revealed the "nature" of light. If Cartwright's interpretation of the presuppositions of Newton's reasoning is correct, it reveals that the concept of a nature plays a key role in some basic types of experimental reasoning. The key results upon which Newton based his inference were that: (i) white light can be refracted by a prism into a spread of differently refrangible rays of many shades of colour, (ii) a prism can *not* subsequently decompose any of the coloured rays further. Now, suppose, as Newton did, that there are only two possible explanations of the observed behavior: either that white light entering the prism is transformed into a variety of coloured rays by the interaction with the prism, or that the white light already contains

---

[37] Ibid., 178.

[38] My caveat about "tendencies" is that it has connotations about frequency of occurrence I want to avoid. Suppose x causes y, 1% of the time that x is present. It seems strange to say that x has a tendency to produce y, but not so strange to say that x has the capacity to produce y.

the various colours, and that the prism just separates them out. Which theory of light is confirmed by the evidence of this experiment?

If you are Wolfgang von Goethe, Newton's critic, you would answer—neither! Goethe believed that there was no reason to reduce the range of possible explanations to just these, because there are many other experiments with light that ought to be considered relevant, and these might suggest other explanations for the observed phenomena. For Newton on the other hand, the range of plausible explanations is implicitly narrowed by the concept of nature. Either it is in the nature of light, or in the nature of a prism that we ascribe the explanation. Since the nature of the prism is not sufficient to account for both facts of the experiment, it must be the facts about the nature of light that account for its behavior. Cartwright's discussion of this point is worth quoting at length:

> Newton focuses on this one special experiment and maintains that the account of the phenomenon in that experiment will pinpoint an explanation that is generalizable. The feature that explains the phenomena in that situation will explain phenomena in other situations; hence he looks to a feature that is part of the inner constitution of light itself. To place it in the *inner* constitution is to cast it not as an observable property characteristic of light but rather as a power that reveals itself, if at all, in appropriately structured circumstances. To describe it as part of light's constitution is to ascribe a kind of permanence to the association: light retains this power across a wide variation in circumstance—indeed, probably so long as it remains light. That is, I maintain, to treat it as an Aristotelian-style nature. That is why Newton, unlike Goethe, can downplay the experimental context. The context is there to elicit the nature of light; it is not an essential ingredient in the ultimate structure of the phenomenon.[39]

If, like Goethe, one believes that all phenomena are locally-realized, situation-dependent products of interactions of "opposites," then there are no intrinsic natures that determine the behaviour of the entities involved in the phenomena. One is then in the position of being unable to determine which experiments carry more evidential weight than others and unable to determine how many experiments are sufficient to establish some general claim as true. But if that is the case, there is no reason to think, as Goethe does, that a diversity of experiments, or more of them, necessarily provides higher quality evidence for any general claim. Newton, on the other hand, is certain that "it is not number of experiments, but weight to be regarded; and where one will do, what need many?"[40] Newton not only thought that in some cases, experiments can provide conclusive demonstrations of the truth of a theoretical claim, but that in those cases, the diversity and quantity of experiments was immaterial!

If Newton was right that his experimental results deductively entailed his conclusions, and rendered all possible objections specious, that could *only* be the case if his experiments revealed some essential feature of light *as such*—some feature that would be present (and invariantly so) from context to context. Newton's lesson is the realist's lesson: if you want to be able to generalize from the results of an experiment, make sure your experiment reveals the context-insensitive aspects of the entities you are investigating—that it reveals the nature of things.

§3. The Composition of Causes Model

The causal realist ought to regard entities as being all embedded in a rich and fluctuating causal nexus in which no entity acts in isolation. Every action that occurs is the actualization of a

---

[39] Cartwright, "Aristotelian Natures," 69.

[40] Issac Newton, "Newton's Optical Papers," in *Issac Newton's Papers and Letters*, ed. I. B. Cohen (Cambridge, MA: Harvard University Press, 1958), 36, quoted in Cartwright, "Aristotelian Natures," 63.

capacity for action in an entity, and every capacity requires the presence of the necessary triggers and enabling conditions for its action and the absence of inhibitors. Even if there is a necessary connection between an entity's capacities and its actions in a given situation, there are no guarantees that nature will comply and give us the right conditions for the exercise of an entity's capacities. Generally, it does not. Smokers do not necessarily die of lung cancer, dropped vases do not necessarily break. There are of course, examples where the consequences of a certain action follow inexorably, insensitive to their surrounding conditions. To use Ducasse's favorite example again, victims of decapitation necessarily die as a result. Obviously these examples are exceedingly rare. [41]

An entity's powers are manifest in certain characteristic modes of activity or characteristic behaviors, modes of activity that are metaphysically hooked up with the intrinsic nature of the thing. The way to determine what the powers of a thing are is to regulate or eliminate all the sources of change in an entity except for one at a time, and see what the entity does under circumstances where each power acts unimpeded. If, as in the normal case, multiple power-actualization processes are occurring simultaneously, they can interfere with one another, even negate one another. For example, the same condition might activate a capacity and a liability that have opposite effects, with a net result of no effect. One way of understanding how multiple, situation-sensitive power-actualizations give rise to actions is to model them after forces in mechanics. Again, I follow Cartwright's argument for the most part.

Cartwright uses the example of the Coulomb potential to illustrate:

> In certain physically impossible contexts (where the bodies involved are massless so gravity plays no role and no other forces are at work either), a charged body causes a force on another of size kq_1q_2/r^2. We also maintain the stronger claim. 'Charged bodies tend to (or 'have the capacity to') exert a force of kq_1q_2/r^2.'[42]

Coulomb's law states that a body O_1 which has a charge of q_1 carries the capacity to exert a force **F** upon another body O_2 (at a distance of r) of charge q_2. Now, even if O_1's Coulomb potential does not produce its characteristic effect on O_2 (because of the presence of gravity, let's say) it would be false to say that the entity *thereby* is said to have lost its power. So long as it is charged, the power remains the same, regardless of whether the power is fully or partially actualized, or not at all.

If we rejected the component power-actualization model of action, how would we interpret the situation-sensitivity of force manifestations? One way would be to reject the assumption and to maintain that every cause generates its full effect every time, independent of the situation.[43] That alternative can easily be discarded. Consider capacities: if a capacity is to be identified with the *action* that is characteristically produced by an entity with a certain nature[44], then we can measure

---

[41] They would be rare, provided they existed at all. Perhaps one could imagine the case where, immediately after decapitation, a band of crazed philosophers pull up with a brain in a vat, and somehow manage to successfully surgically reattach this brain to our poor beheaded friend to keep his autonomic nervous system functioning for a few minutes. The body then has a heart attack and dies. The cause of death? Not decapitation, but cardiac arrest. This typical (in philosophy of mind circles, anyway) kind of counter-example would show that it is not necessarily true that all victims of decapitation die as a result.

[42] Nancy Cartwright, "Reply to Eells, Humphreys and Morrison," *Philosophy and Phenomenological Research* LV, no. 1 (1995): 180.

[43] The other possible rejoinder would be to maintain that there really is no motion, but I am assuming that Parmenideanism is out.

[44] I am assuming that being charged is an intrinsic property of the particles that have it, which is due to some further level of fine structure to which we have no access at present.

the extent of a capacity-actualization by measuring the action it produces. If the capacity in question is the capacity to exert a force, then it is measured by the acceleration that the capacity produces. The acceleration of O_2 *is* the actualization of a capacity of O_1. If O_2's action is countered by an equal and opposite force from some other entity (or its associate field) then O_2 experiences *no* net acceleration. What sense would it make to say that the effect of O_1's power *and* that of the counteracting force are both actualized? Commenting on a similar proposal suggested by Mill, Cartwright says,

> To claim that a motion exists in a body even when that body is at a standstill, and passes all the conventional tests for being at a standstill, is to forsake empiricism, and to do so in a way that violates its fundamental tenets … for it severs the existence of the motion from all our standard methods of measuring motions.[45]

In order to comply with the requirements of empiricist methodology, Cartwright maintains that we need to keep existence claims for capacities supported by their measurability. A power in general is to be identified with the *action* that is characteristically produced by an entity with a certain nature, so the condition for the measurability of the capacity ought really be regarded as the condition for the measurability of the capacity-actualization. Recall that according to the power-actualization process theory of causality, entities are causes, while actions, or capacity-actualizations, are the effects. Therefore, the measurement of a capacity is the measurement of the *effect*, not the cause.

Cartwright's suggestion is that we should interpret the composition of forces as the composition of causings (as compositions of power-actualizations) as an argument from analogy. The formal similarities between the resultant motion of a body due to composition of forces and the resultant action of an entity due to the composition of causes, establish that the analogy of powers to forces is a good one. We can use the model to explain the peculiarities of power-manifestation such as their variable and partial realization from one situation to the next in terms of their sensitivity to the presence and absence of the appropriate conditions. The analogy shows how the Coulomb potential may be interpreted as a capacity, but *any* potential could have been used. The validity of the concept of capacity is generalizable to other contexts where forces are at work. Since forces are at work everywhere, the potential domain of validity of an ontology including capacities may well extend to all of physics.

§4. Explaining Mathematical Practice in Social Science

In economics, the equation for the demand for some class of goods is expressed by an equation of the form:

(*D*) $q = \alpha p + u$

where q is quantity, p is price, α is demand elasticity and u is an error term. The quantity appearing on the left side of the equation indicates its theoretically presumed causal dependence on price. The demand function is supposed to model the quantitative response of preferences to changes in price, but not just with respect to some particular set of data. The value of modeling only accrues to economists when the models capture some behavior that is invariant with respect to changes in context. Econometricians assume that price has a stable tendency to influence demand, and that the tendency has a measurable strength, represented by the fixed parameter α. The two features of *D* of interest here are its being a *first-order* function of p, and the *invariance* of its parameters. The constraints imposed by the postulates of parameter invariance and linearity make possible the

[45] Cartwright, *Nature's Capacities*, 179.

confirmation of the demand function for some class of goods.[46] While the postulate of linearity seems uncontroversial, it is parameter invariance that does most of the work in the argument for capacities. Again, Cartwright argues in an **IBE**-style: if capacities are at work, then they will be stable across all relevant contexts. If capacities are at work and stable across all relevant contexts, then that ought to show up in parameter invariance in the econometric models. Thus, the invariance of parameters that we do observe is best explained by the naturally stable activity of capacities:

> it is just the feature of stability that I have pointed to in econometric structures that must be presupposed in any discipline if we are to infer from the statistics we observe in one set of conditions to the effects that might be achieved by varying these conditions.[47]

If probabilities are a reliable guide to causes, it can only be because context-invariant stable capacities that justify inferences from one context to features of another are presupposed by the methods themselves.

This argument leaves itself open to one fairly serious objection, voiced by James Woodward:

> It is simply an empirical fact that such parameters are often not stable from situation to situation in the sorts of contexts in which econometric techniques are used. Thus, if capacities must satisfy the sort of stability requirement that Cartwright favours, econometrics discovers few, if any, facts about capacities.[48]

Of course in complex systems, the parameters characterizing modes of the aggregate are only epiphenomena—they themselves are not causal but only the individual decision-making units in the economy are.

§5. Explaining Why Probability Increases Recur

It is a basic fact of human psychology that people tend to view certain kinds of correlated events and/or properties as causally connected. Some people have "lucky hats," which they think will increase their future chances of good fortune, others take Vitamin C, thinking that it will decrease their risk of developing a cold. Whether spurious or systematic, there is an expectation that observed regular connections will continue to be repeated—that they are not "coincidences." Take the semi-plausible Vitamin C case. When it is presumed that the long-run chances of getting a cold are, overall, less if one takes Vitamin C than if one does not, that is assuming that Vitamin C prevents (is a negative causal factor for) colds. In other words, the cause, C, decreases the probability of the effect, or $P(E/C) < P(E/\sim C)$. In general, genuine causal factors are supposed to alter the probability (frequency) of the effect. Typically what is done to test causal hypotheses is to get samples of two groups, give everyone in one group the alleged causal factor, and prevent anyone in a second group from getting it. The difference in the frequency of the effect between the two groups over some fixed interval of time is then determined. This differential measures the strength of the causal factor—the strength of the capacity of the cause to produce the effect. Now, it is possible that taking other vitamins can help prevent colds, but so can drinking lots of water, eating a nutritious

---

[46] The logic of confirmation here is supposed to work by bootstrapping: first, we use background economic theory to deduce the form of the structural model. Then we use the econometric data to estimate the parameters of the model. The data combined with deductions from background theory together *imply* the specific quantitative theory. See Cartwright, *Nature's Capacities* for details.

[47] Cartwright, *Nature's Capacities*,158.

[48] Woodward, "Capacites and Invariance," 325 n. 18.

diet of foods, regularly washing one's hands, *etc.* The rate of occurrence of any other conceivably causally relevant factor must be the same between the two groups—the profiles of the people in the two groups much match as closely as possible. In order to test a causal hypothesis, we need a causal criterion that addresses the need for matching—that in the control and experimental groups each person either has or does not have some trait, habit, *etc.* Cartwright's "causal criterion" (CC) suggests one way of formulating it. I suggest a more general formulation which adds (i) an allowance for the long and variable temporal lags between causes and effects that are possible, (ii) negative causal factors or preventatives, and (iii) a measurement ε, of the strength of the capacity:

(***CC+***) C causes $E \equiv_{df}$

$$\lim_{\Delta t \to \infty} P(E_t | C_{t-\Delta t} \pm F_1 \pm F_2 \pm ... \pm F_n) - P(E_t | \sim C_{t-\Delta t} \pm F_1 \pm F_2 \pm ... \pm F_n) \approx \varepsilon$$

where $\varepsilon \neq 0$ and $\{F_1, F_2, \ldots, F_n, C\}$ is a complete causal set for E, *i.e.*, any and all causes of E are included. ε is the value of the strength of the capacity of C to cause E if $\varepsilon > 0$ or prevent E if $\varepsilon < 0$. This condition says that the limit of the difference must be about ε in any population P_n homogeneous with respect to the Fs. In other words, the probability differential must appear between any two appropriately matched sample groups drawn from any population.

If the probability differential converges to ε for large values of t-Δt in every homogeneous population P_n, then the cause C displays "contextual unanimity." We are now in a position to make another argument for why causality out to be interpreted as a power-actualization process. Again, I follow Cartwright's lead. She reasons as follows:

1. If C causes E in P, then C has the capacity to cause E in P
2. If C has the capacity to cause E in P, then C has the capacity to cause E in any P (C causes E, *simpliciter*)

(3.) If C has the capacity to cause E in any P, then C has the capacity to cause E in all Ps (contextual unanimity)

Cartwright concludes, "to believe in contextual unanimity is to believe in capacities, or at least is a good way along the road to that belief."[49] But the conclusion does not follow, at least not deductively. The most Cartwright is entitled to claim is that the reality of capacities is the best explanation for how contextual unanimity is possible—for why increases in probabilities recur throughout the class of populations homogeneous with respect to the background causal factors for E.

If it is true that C causes E, *not* in just some appropriately homogeneous populations, but in *all* such populations, then a universal claim is being made—it is no longer about the token Cs in specific populations, but is a claim about Cs *qua* type. In other words, when contextual unanimity holds, we are seeing something about the capacity of Cs to cause E in virtue of being the kind of entities that they are—that they are acting according to their nature. This is why the reality of capacities is, in fact *presupposed* by probabilistic approaches to causal inference.[50]

---

[49] Cartwright, *Nature's Capacities*, 145.

[50] I say "probabilistic approaches to causal inference" rather than "probabilistic theories of causality" for a couple of reasons. First, the so-called probabilistic causality theories are focused on the epistemology of causal hypothesis confirmation—when can you say that C causes E? They do not address themselves to the question of what causality *is*, just to questions concerning when you are entitled to say you've found it. Second (and this is where I disagree with Cartwright) causes do *not* act probabilistically in this sense: metaphysically, there is no such thing as chance. Probability and chance are concepts applicable to contexts in

§6. Explaining the Analytic Method

Cartwright argues that distinctive features of experimental practice and scientific methodology could not adequately be explained if there were no "natures." The existence of natures is a methodological presupposition of empirical inquiry, from experimental design to deducing general laws from phenomena. Her basic point is this: that "one cannot make sense of [these aspects of] modern experimental method unless one assumes that laws are basically about natures."[51] She argues for the following conclusion as a way of explaining why an ontological commitment to natures is implicit in the experimental method:

> Situations that lend themselves to generalization are very special, and it is these special kinds of situation that we aim to create, both in our experiments and in our technology. My central thesis here is that what makes these situations special is that they are situations that permit a stable display of the nature of the process under study, or the stable display of the interaction of several different natures.[52]

It is only if an experiment succeeds in selectively revealing the operation of natures that we can infer a general law about the entity or process whose behavior is expressed by the law:

> we need the notion of natures to … infer from the result of the experiment some general laws that describe what happens, just in this experimental situation, whenever the experiment is run again.[53]

The account she is offering is realistic about natures, and thus is explicitly anti-Humean: "The facts about an experiment that make that experiment generalizable are not facts that exist in a purely Humean world."[54] To further polarize the opposition between her brand of realism and Humean empiricism, she identifies the kinds of experimental circumstances in which one is "entitled to generalize from even a single case."[55]

The primary burden of argument for Cartwright is justifying the twin claims that science presupposes natures and that science works *because* it presupposes natures. She argues by example; in the first example, decomposition of forces as a way of isolating and identifying the nature of electrostatic attraction is discussed. This shows how experimentation presupposes natures. The second details the planned Stanford Gravity-Probe-B experiment designed to reveal the natures of relativistic frame-dragging and the geodetic effect by measuring whatever precession is induced by the Earth's gravity on spinning gyroscopes. It shows how scientific inference presupposes natures.

Consider the first example. We know that in physics, the accelerations of particles are caused by the forces that act on them. Generally, several forces are at work on a particle simultaneously; the overall motion of a particle is the vector resultant of a continually changing flux of forces involved. There are no Humean regularities in nature as far as the actions of particles are concerned. A particle, at any given spatio-temporal location, is subject to a unique combination of forces; the conditions are probably never exactly repeated. How, then, are we able to explain or predict what a particle will do when it is introduced into a region full of particles of various kinds and overlapping

---

which information is incomplete, by accident or design, and are epistemic concepts through and through. My view imposes certain constraints on the interpretation of quantum probabilities—it is consistent with a non-local realistic hidden variables theory, but not with the Copenhagen interpretation.

[51] Cartwright, "Aristotelian Natures," 47

[52] Ibid., 52.

[53] Ibid., 56.

[54] Ibid., 53.

[55] Ibid., 55.

fields of different types of force? According to the analytic method,

> to understand what happens in the world, we take things apart into their fundamental pieces; to control a situation we reassemble the pieces, we reorder them so they will work together to make things happen as we will. You carry the pieces from place to place, assembling them together in new ways and new contexts. But you always assume that they will try to behave in new arrangements as they have tried to behave in others.[56]

In order to predict the behavior of some type of particle which is typically subject to various mixes of forces over time, it is necessary to isolate how that kind of particle acts when it is subject to each of the forces one at a time. In order to accomplish this, we need to construct a situation in which we can measure the force between two charged bodies, when no other forces are at work, another situation in which we can measure the gravitational attraction between two massive bodies when no other forces are at work, and so on. Why? We are interested in what the bodies do "on their own," that is, in virtue of being charged, massive, *etc.*

Consider how electrostatic and gravitational forces combine to determine the action of an object. If two particles, one with electrostatic charge q_1 coulombs and the other q_2 coulombs, are placed a distance r apart from one another, they will attract each other with an electrostatic force given by:

$$F = 4 \pi e_o \, q_1 q_2 / r^2$$

In this equation, e_o is a constant: regardless of the situation, the force is always the same. This equation does not tell us what forces charged particles regularly (if ever) experience, but it tells us what it is in the nature of a particle to experience, insofar as it bears a charge. A similar situation holds if same two particles have mass **M1** and **M2** respectively; they will attract each other with a gravitational force given by:

$$F = G \, m_1 m_2 / r^2$$

For particles such as electrons that are massive as well as charged, the forces they actually experience will be a function of what they experience *qua* charged, and what they experience *qua* massive. Outside of a contrived, controlled environment, in which the effects of all the other forces are identified and either fixed or reduced to near zero, the "laws" of the two equations above do not report the actual force experienced by the particles in question. It is usually only in artificially constructed circumstances that we can isolate one aspect of an entity's nature and observe the dispositions it has in virtue of that aspect of its nature.

Compliance with the analytic method involves analyzing complex, non-uniform behavior into simple, stable behavior by separating out the individual natures that jointly contribute to the production of complex phenomena. The analytic method works. Generally we *can* explain and successfully predict the behavior of things in more complex situations. Developing technology would be impossible otherwise. But that success, and the expectation of future success is only justified if things "act in *accordance with their nature.*"[57] The assumption is, of course, that an entity is something that has a nature—it is what it is and it acts accordingly.

This assumption not only makes sense of experimental designs which often go to incredible trouble just to see single aspects of an entity's nature, but it also is a premise needed to justify inferences we make to general propositions describing the nature of a thing or process. Cartwright

---

[56] Ibid., 49.

[57] Ibid., emphasis added.

explains:

> we measure, successfully we think, the charge or mass of an electron in a given experiment. Now we think we know the charge or mass of an electron in a given experiment. Now we think we know the charge or mass of all electrons; we need not go on, measuring hundreds of thousands. In so doing, we are making what looks to be a kind of essentialist assumption: the charge or mass of a fundamental particle is not a variable quantity but is characteristic of the particle so long as it continues to be the particle it is.[58]

It is only when we can safely say that we are seeing the action of an entity on the basis of a single aspect of its nature that we are justified in maintaining that the measurements we make of its qualities under those conditions are the data from which we may infer a "law" of its behavior.

Another illustration of the need for natures that Cartwright presents involves the composition of forces. She explains that for two charged particles, to say that "it is in their nature" to experience a force proportional to q_1q_2/r^2 is to maintain that "they *can* experience this force if only the right conditions occur for the power to exercise itself." Alternatively, it means that "their tendency to experience it persists, even when conditions are not right." A more perspicuous formulation of the point is this: a particle that has the property "bearing charge q_1" has the liability to experience a force of kq_1q_2/r^2 upon it in the presence of another particle of charge q_2 at a distance r from it. (It also has the capacity to attract or repel q_2 with a force kq_1q_2/r^2.) A liability is a tendency or disposition to be affected by the action of some agent. It is true that an entity with a certain liability has, "in its nature," the power to respond to the action of another, but a liability is not "something like" a nature. Natures, as Cartwright agrees, are not irreducible subsisting essences but intrinsic structural features of entities.[59] They are, as Aristotle maintained, "causes and principles of change" that give rise to characteristic modes of behavior, and those modes of behavior will be necessarily connected to those structures, so long as the structure is what it is.

The second example, involving the Gravity Probe, highlights the kind of "totally controlled experiment"[60] it takes to create an environment in which some phenomenon can be identified, its nature established, and some kind of law derived. The relativistic precessions of a gyroscope predicted by Leonard Schiff are two incredibly subtle effects, but if they are observed to occur as predicted, the phenomena would be confirming evidence for the theory of general relativity. Just as in the case of charged particles, where we can only identify the nature of the electrostatic interaction in the case where the masses "go to zero," in the case of relativistic coupling, we can only identify the relativistic precession if the gyroscopes are virtually perfectly spherical and perfectly homogeneous, the temperature is near absolute zero, the apparatus is frictionless and in a super high vacuum, and the axis of rotation of the gyroscopes is perfectly aligned to a reference star. Any compromise of the integrity of the design would risk ruining the experiment. The frame-dragging and the geodetic effects would be swamped by "noise" caused by other uncontrolled factors which could not precisely be calculated away.

If the Gravity Probe experiment is successful, a high-level generalization like this could be inferred:

(H) The coupling of a spinning gyroscope to curved space-time is described by law *L*.

---

[58] Ibid., 51.

[59] Cartwright, "Aristotelian Natures," 47.

[60] Cartwright, *Nature's Capacities*, 66.

Also, a low-level generalization is warranted:

(L) Any spherical fused-quartz gyroscope surrounded by a molecular layer of superconducting niobium, suspended electromagnetically in a 10^{-11} torr vacuum, cooled to 1.8 K by liquid-helium injection, spinning in deep space at 10,000 rpm, will precess at a certain drift rate rate n_s.

This will allow the calculation of the geodetic rate of precession, Ω_G, resulting from the motion of the gyroscope due to the curved space-time around the earth, and the motional rate, Ω_M, resulting from the rotation of space-time around the earth dragged by the earth's own rotation.[61]

While the experiment will generate some particular value for the two rates, would the conclusion (L) be justified on the basis of a single experiment? Would not repetition of the experiment with a variety of different configurations be necessary in order to secure confidence in the result? Even if (L) could be justified by confirming it with a variety of experiments, how could (H) be inferred? According to Cartwright, the reality of natures licenses the inference from the experiment to (H), *and then* to (L):

> It is just to the extent that we acknowledge that the experiment is well designed to find out the natures of the interaction, described in the higher-level law, that we are entitled to accept the low-level generalization on the basis of the experimental results.[62]

She admits that the ontological underpinning that supports these claims implies a "radical shift from the picture in which the conventional view of laws is embedded…."[63] If natures exist, and an entity acts in accordance with its nature, so long as it continues to be the thing that it is, then repetition of an experiment is epistemologically redundant. The same thing will act in the same way under the same conditions.

§7. The Aim of Science

In addition to Cartwright-inspired arguments, here is one in a similar vein from Rom Harré, who maintains that

> the aim of science is to try to find the structures, states and inner constitutions from which the phenomena of nature flow. … It is, in short, to look for the causal mechanisms of which the patterns and regularities of phenomena are the effects.[64]

Could the aim of science be, on the contrary, the discovery of theories which save the phenomena—which latch onto the regularities in nature? This is "simply a mistake, arising from preconceptions of logic and from ignorance of scientific practice."[65] Harré appears to be advocating what van Fraassen refers to as "traditional" philosophy of science in which

> everything is subordinated to the aim of knowing the structure of the world. The central activity is therefore the construction of theories that describe this structure. Experiments are then designed to test these theories, to see if they should be admitted to the office of truth-bearers, contributing to our world picture.[66]

Harré and Madden subscribe to a version of convergent realism, whereby scientific progress

---

[61] Ibid.

[62] Cartwright, "Aristotelian Natures," 55.

[63] Ibid.

[64] Rom Harré, *The Principles of Scientific Thinking* (Oxford: Oxford University Press, 1968), 102.

[65] Ibid.

[66] Bas C. vanFraassen, *The Scientific Image* (Oxford: Clarendon Press, 1980), 73.

amounts to a gradual, progressive unveiling of the real essence of things by empirical inquiry.[67] The "real essence" of a thing or substance is only revealed when the analysis of it is "complete," whatever that means. The notion inherits much of the obscurity of its Lockean heritage. The "nominal essence" that plays the role of a "stand-in" for the real essence, which we use in the empirical identification of a thing, is one that involves sensible qualities identifiable without special techniques of experiment or observation. Harré and Madden maintain that certain kinds of discoveries count as evidence in favour of thinking that at least an approximation to the real essence has been achieved. A more perspicuous way of making the same point might be to say that certain kinds of discoveries count as evidence in favour of thinking that a more accurate—more truth-like—identification of the nature of a particular or kind of substance has been made. For example, the isolation of oxygen as a gas that is necessary for combustion, and is consumed by oxidation, gave a more accurate picture of the *nature* of combustion. The identification of heat as the mean kinetic energy of molecules in motion in (nearly) empty space supplanted the theory of caloric as an indestructible, highly elastic, self-repellent, all-pervading, massless, "imponderable fluid" which explained the production and transfer of heat. The identification of heat with motion identified accurately the *nature* of heat.

A corollary of the idea that science aims to discover the intrinsic nature of things is that discovering the nature of things with greater accuracy is necessary to give meaning to the idea of scientific progress.

4.4 The Logic of the Arguments

The fundamental grounds for the introduction of the concept of nature into our ontology is the "object persistence" argument. If the changes in a thing are not related to some persistent feature of the entity involved, then there is no thing that is the subject of the changes—in effect there are just changes, without a *changer*, which is an incoherent notion.

The argument for the ontology of powers is somewhat more involved, however. The overall thrust of the arguments of §4.3 is to show that methods of science cannot be adequately explained without the ideas of power and nature. The arguments presented were selected specifically because they each conclude that a that a philosophical analysis of various aspects of causal inquiry in science requires these concepts.

The conclusions Cartwright makes are two: first, that inferences to general laws based on the results of highly controlled experiments designed to get close to the "ideal" case are valid, and that they are valid because they are designed to isolate and characterize specific natures; second, that "in order to understand what makes experiments special, what ensures that we can generalize from them, we must employ concepts repugnant to a Humean, such as nature, power, impediment …."[68]

Abstracting from the particulars of the arguments she offers for the extended ontology of realism, a basic argument style emerges. Drawn out in premise-conclusion form it goes something like this:

(i) Method X is an effective tool for performing some aspect of scientific activity (*i.e.*, experimentation, drawing conclusions, formulating explanations).

---

[67] For a classic and provocative critique of convergent realism, see Larry Laudan, "A Confutation of Convergent Realism," *Philosophy of Science* 48, no. 1 (1981): 19-49.

[68] Cartwright, Nature's Capacities, 56

(ii) Method X makes presuppositions about the existence of Y (i.e., natures, capacities, unobservable entities)

(iii) The effectiveness of X cannot be adequately explained unless the presuppositions about the existence of Y that X makes are correct.

(iv) It is rational to accept the existence of that which provides the only adequate (or the best) explanation for the *explanandum* in question.

∴ It is rational to accept the existence of Y.

Recently,[69] Cartwright has qualified the conclusion somewhat. She maintains that there are indeed natures in our "world picture," meaning that as a *concept of method*, "nature" is indispensable for understanding how and why science works. Science works *as if* there were natures in the world; it is therefore a rationally effective strategy to approach scientific experimentation and empirical inference on the assumption that natures are real.

The logic of the argument resembles inference to the best explanation. If science is best interpreted as an enterprise that takes the existence of natures for granted as a methodological posit, then that is at best *inductive* evidence that the implicit ontological claim is true. Despite the fact that Cartwright regards inference to the best explanation as not capable of establishing the truth of claims in general, one variant of it *is* endorsed—inference to the most probable *cause*. When the most probable cause is not just marginally more likely than not, given the evidence, but stands as the only basis for a causal explanation of some fact, then the argument bears a resemblance to a "transcendental" style of argument. In this section, I want to try to identify whether such arguments are valid.

§1. Validating Ontological Concepts

We need to distinguish between two closely related kinds of claims: (a) claims affirming the existence of an entity of a certain type, (b) claims for the validity of a concept designating a type. This distinction will play an important role in assessing the logic of the extension of the "core" ontology of realism to include the types of things required for an adequate general theory of causality. One particularly germane example is the claim "capacities are real" (or "capacities exist"). This claim differs in subtle but important ways from the claim " 'capacity' is a valid concept." The kinds of evidence and argument needed to justify the two claims are different.

Consider, as an example of the first type, the so-called "Ontological Argument" for the existence of God. As the paradigmatic example of a class of argument, ontological arguments are generally taken to be "existence proofs."[70] Now, if the claim that "capacities are real" was interpreted to be a claim similar to this, the claim would be open to misinterpretation. Capacities are not like gods, let alone electrons, quasars, cats and dogs. To use one of Cartwright's favorite examples, consider the claim "aspirins carry the capacity to relieve headaches." This does not mean that an aspirin has some ghostly, occult, discrete bundle of capacities lurking within it. It means that an aspirin acts in a special way, given certain conditions, in virtue of its particular structure. The claim "capacities are real" presupposes a more generalized conception of reality than that of entity-

[69] Cartwright, "Aristotelian Natures."

[70] Of course, it was a notoriously *unsuccessful* type of ontological argument. The fundamental error of that style of argument is the premise that we can count on reality to contain whatever a process of consciousness decides "has" to exist in reality. It is invalid to reason from the necessary truth of some existential proposition arrived at *a priori*, to the actual satisfaction of that proposition in the world. In short: No existence in, no existence out.

existence; it represents the more generalized commitment to the reality of a *type*. Now, the reality of a type is simply the non-emptiness of the category of existents associated with the type. It does not imply the validity of the concept used to refer to the category. The validity of a concept entails, but is not entailed by, the existence of instances of the concept. The generalized conception of ontological commitment that I propose to adopt is associated with the validity of concepts—one in which the veracity of an existence claim for an entity of a given type is a necessary but not a sufficient condition for the validity of the concept of that type.

I assume that concepts are the products of cognitive processes—they reflect the outcome of our attempts to organize and systematize our knowledge of the world. As such, a concept is never an island of knowledge; knowledge is represented by relations (esp. predicative relations) among concepts within a dynamic, complex multi-dimensional network. The introduction of new concepts into a system of knowledge ought to increase the amount of knowledge represented within it, so arguments for the introduction of a new concept must show what *work* the concept does for the knower. Sometimes concepts can be added piecemeal, sometimes a large number of concepts are supplanted by revised ones. In order to establish whether the introduction of a new concept is warranted, I propose that we require an affirmative answer to the question of whether there is anything that we observe in reality that gives rise to a need for it. This criterion implies two conditions for conceptual validity which can then be used to formulate criteria of adequacy for concept-validation arguments. They are:

1. *Referentiality*. The validity of a concept entails that the concept is a representation of something real—that the concept must be *genuinely referring*, or have some content based on perception or observation. Since perceptibility is proof of existence, the reducibility of the content of a concept to its perceptual-level referents implies that the ontological suppositions of the concept are legitimate.

2. *Epistemic Utility*.[71] Concepts ought to serve our epistemological practices—to help us make distinctions, identify relationships, isolate common features, and unify observations. But in order to facilitate these practices, we need concepts that satisfy certain requirements. Among these are:

   (i) Coherence: A concept is a mental representation of a category of particulars. Categories formed arbitrarily will generally be useless. The categories themselves need to comprise existents that "belong together"—that have some dimension of similarity in virtue of which the members of the category form a "natural" kind. The need for conceptual coherence—the need for concepts that refer to objectively homogenous (in some respect) classes of things reflects the need for naturalness of categories.

   (ii) Economy: To say that two or more existents are both instances of the same concept is to regard them, for some purposes, as identical. Cognitive economy is a product of the process of abstraction—*i.e.*, the omission of individuating characteristics among the members of a category and their subsequent integration into a single mental unit. Concepts reduce the amount of information necessary to retain about particulars. To remember and treat everything in one's environment as different would require vastly more cognitive capacity that we have, both in memory capacity and retrieval response speed.

[71] This account draws from those of Lloyd K. Komatsu, "Recent Views of Conceptual Structure," *Psychological Bulletin* 112, no. 3 (1992): 500-26, and Kelley, "Abstraction."

(iii) Informativeness: Given that the referents of a concept are not identical, abstraction inevitably leads to a loss of information about the members of the category it represents. Minimization of information loss is to be achieved by retaining the *essential* features of the particulars whose individuating characteristics are disregarded—that is, those features whose common presence in the particulars is associated with the greatest number of *other* features had by the members of the category represented.

Together, these entail that a valid concept has some stable, well-defined meaning.

Examples of invalid concepts might be "caloric" and "right-wing," as they exemplify concepts that fail the referentiality and epistemic utility criteria, respectively. Caloric was the pre-thermodynamic entity supposedly responsible for heat. While the concept appeared to account for some phenomena, and thus had epistemic utility for a time, it had no referentiality. Heat turned out to be a property of bodies, not a body itself, and so the concept *caloric* was a category mistake; an invalid concept. "Right-wing" is a term with a slightly pejorative connotation, employed almost entirely in rhetorical contexts to manipulate political opinion. It fails the condition of coherence, since the term is applied to a hugely divergent collection of political ideologies, from the advocacy of the totalitarian state (*e.g.*, fascism) to the virtual absence of the state (e.g., anarcho-capitalism). It may be applied to those seeking radical change, or those wanting to maintain the status-quo. It is uninformative, since it has no stable meaning that would permit its use in inference or explanation. The question of who is or is not a member of the class "right-wing" is unanswerable.

The above examples indicate the kinds of considerations that are relevant to establishing the failure of a particular concept to meet the conditions of validity for concepts in general. The criteria of adequacy of an argument for the validity of a concept are implied straightforwardly by the conditions for the validity of concepts. An argument for the validity of a concept is a good argument if it successfully establishes that the concept satisfies the conditions of referentiality and epistemic utility. In addition, it must satisfy the general plausibility requirement for arguments in general.[72]

While referentiality is a necessary condition for conceptual validity, it is not a sufficient condition. For example, the existence of the referents of the concept "electron" is a necessary condition of the validity of the concept "electron." But it is not sufficient—one could form the concept "glyques" consisting of the referents of the heterogeneous set {electrons, bagels, scallops} all of whose referents exist. The concept "glyques" would be invalid, not because the reference class is null but because it fails to satisfy the coherence condition, it is of no epistemic utility. Ontological commitments must be supported by arguments that demonstrate both the epistemic utility and the referentiality of the concept in whose favour it is making a case.

§2. The Epistemic Utility of Power Concepts

The methodological arguments for capacities and natures do not establish referentiality, even if the conclusions are worded as if they did. What one is logically entitled to claim, based on a methodological argument, is the epistemic utility of the concept. And Cartwright's arguments satisfy a pattern of methodological argument that I propose to call a "conceptual indispensability argument."

---

[72] The limit of argumentative plausibility is deductive validity. The premises of the argument must be consistent with what we know to be true, and if the truth of the premises be granted, the falsehood of the conclusion must be highly unlikely. A plausible argument need not deductively entail its conclusion nor render its conclusion incorrigible or indefeasible. A plausible argument just has to provide grounds, based on the available evidence, for regarding the conclusion of the argument as sufficiently likely to be true.

The arguments of the previous two sections welcomed powers and natures into our ontology by showing their indispensability for understanding the basic inferential and explanatory practices of science. A concept is indispensable if it cannot be eliminated or replaced in favour of others without incurring epistemic losses. Valid instances of conceptual indispensability arguments identify the presuppositions or bases of some other concept, theory or practice. These arguments exploit the relations of conceptual and epistemic dependency implicit in the view that conceptual systems exhibit a hierarchical structure. What is the logic of those arguments?

The simplest kind of conceptual indispensability argument is one that exploits self-refutation. It is also the strongest, logically. Suppose one encountered a radical behaviorist or eliminativist who denied that consciousness existed. Whether consciousness is an illusion, or a self-deception, the bottom line is that we do not have minds, and that the concept of consciousness is meaningless. All that is sufficient to refute these positions is to point out, in Cartesian fashion, that the very *act of denial* (not to mention affirmation) is a phenomenon of consciousness; the reality of consciousness is presupposed by any act of thought, including any attempt to invalidate the concept of consciousness. Since all attempts to deny the existence of consciousness are necessarily self-refuting, this shows that the concept is inescapable. From this, we might say that the concept is universally indispensable.

A slightly more complicated case of conceptual indispensability is provided when the indispensability is relative to a specific context. For example, to understand the theory and methods of differential calculus, one must first understand the idea of a ratio. The derivative is defined as the limit of a ratio:

$$f'(x) = \lim_{x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

the slope of the tangent line to f(x) at x, and (for single-valued functions) represents the rate of change of the quantity represented by the dependent variable relative to that of the independent variable. Is it possible to explain the concept of the derivative without using ratios? If not, then one cannot understand the concept without a prior knowledge of ratios. In other words, the concept of the derivative necessarily depends on the concept of the ratio. Let us call arguments for the validity of a concept which depend on necessary conceptual dependency considerations *epistemic indispensability arguments,* to capture the idea that the meaning of some concepts cannot be known until one has understood the meanings of the concepts in terms of which they are defined.

Now let us change the example slightly. Does the conceptual dependence of the derivative on *limits* imply that differential calculus cannot be explained without using *them*? Here the situation is slightly more complicated, since the theorems of differential calculus were originally formulated by Newton using infinitesimals instead of limits. The concept of a derivative was fully intelligible and differential calculus of considerable utility in science and engineering for a century and a half before Cauchy's reformulation of the basic definitions and theorems of analysis in terms of limits. This shows that there is no *necessary* dependency of stating and solving differential equations upon the concept of a limit. While Newton's differential calculus was serviceable as a tool for calculation, and facilitated the solving of practical scientific problems, there were certain problems that could not be resolved because there are *certain forms of inference that cannot not be expressed* unless the concept of the derivative is defined using limits. This second example represents a different kind of indispensability—one that characterizes the dependency of solutions to specific mathematical problems upon the limit-formulation of the derivative. In particular, metatheoretic results such as existence and uniqueness proofs are proofs whose statement required the level of logical rigour

introduced into the calculus by the theory of limits. Such proofs are of utility to mathematicians, because they allow investigation of the general properties of problems. In the theory of differential equations, for example, limit-based metatheory allows one to establish whether a problem is "well-posed." A "well-posed" problem is one for which there is a proof that it has a solution, has a unique solution, and depends in a reasonable way on the data—*i.e.*, on the empirical magnitudes that the parameters of the equations involved may take. By being able to express proofs for the solubility or the insolubility of specific classes of problems, mathematicians are able to define more clearly where valid areas of further research lie.

The metatheoretic proofs for the existence and uniqueness of solutions to problems unattainable without the theory of the limit are what validates the limit concept. In other words, existence and uniqueness proofs *presuppose* the existence and definability of limits. By this fact, limits are introduced into the "ontology" of mathematical objects. This way of showing the indispensability of the concept can, I think, be generalized to the notion of an *inferential indispensability argument* for the validity of a concept. The general form of this type of argument is one whose conclusion is a statement regarding the validity of a concept and whose premises gather evidence of arguments whose validity presupposes the truth of the propositions in which those concepts appear, and *eo ipso*, the validity of the concepts necessary to express those propositions. More formally, the form of the argument is this:

(1) Let C be a conclusion which is taken to be justified by some set of evidence-statements $\{p_1, p_2, \ldots, p_n\}$, but not deductively entailed by them.

(2) If there exists another statement q such that $\{p_1, p_2, \ldots, p_n\} \cup q \vdash C$, and for any statement r $(q \neq r)$ such that $\{p_1, p_2, \ldots, p_n\} \cup r \nvdash C$, then q is true.

(3) For any true statement $q \equiv \pm S \pm P$[73], where S and P are valid concepts.

$\therefore$ S and P are valid concepts.

The following postulate states a criterion of conceptual validity based on the availability of such arguments:

IIC *The Inferential indispensability criterion*: If a pattern of argument whose conclusion is (so far as we know) never false when its premises are true requires, in order to demonstrate its logical validity, the truth of propositions containing concepts which cannot be eliminated in favour of others without a loss of logical strength or of intelligibility, then those concepts are valid, provided that they are defined in a way that preserves the soundness of the inference in which they appear.

When this criterion is applied to an argument, it licenses the inference to the truth of the "gap-filling" propositions (q)—those required in order to explicate or "rigorize" the logic of the argument from $\{p_n\}$ to C. It also validates the concepts necessary to express the gap-filling premises. Of course, figuring out *which* premises are necessary to justify the inferences in question is no trivial matter, given the stringency of (2) in the above argument. One way of doing this is to use a form of eliminative induction, such as "inference to the best explanation" (**IBE**).

Realists in the philosophy of science have often used (or been accused of using!) **IBE** as a means of arguing for their positions. The argument maintains that if the (approximate) truth of science's theories best explains their success, and that success is evident, then it is reasonable to infer the best explanation: the (approximate) truth of science's theories. The scientific realist interprets the success of a theory as meaning that the theory gets the structure of a bit of reality

[73] This is the generalized normal form of a proposition in Fred Sommers' term logic.

pretty much right. The simplest version of **IBE** is:

(**IBE)** Theory T is (approximately) true → T is successful

T is successful

∴ T is (more probably approximately) true[74]

When the anti-realist argues against **IBE**, the argument usually centers on the truth of the conditional premise in the argument rather than on the logical validity of the argument form, which the anti-realist cannot avoid also employing, as in this example:

(AR) Theory T is empirically adequate → Observation O

Observation O

∴ Theory T is (more probably) empirically adequate

Cartwright's argument above can be represented, albeit somewhat oversimplified, as an elaboration of this variant of **IBE**:

> The truth of the presuppositions scientific method X makes about the existence of Y best explains why scientific method X works in practice.
>
> Scientific method works in practice.
>
> ∴ The presuppositions scientific method X makes about the existence of Y are true.

Cartwright's employment of **IBE** is in fact somewhat more narrow than is usual, since she infers just to the existence of some explanatory entity without requiring that the characterization of it be accurately given in any particular theory. She rejects **IBE** as a general pattern of reasoning, endorsing only a particular form of it—what she calls "inference to the most probable cause." Perhaps the most interesting aspect of the arguments that take this form is the ontological conclusions that the inferences license if they possess sufficient logical strength. They analyse scientific methods and practices for whatever ontological presuppositions are made, and argue from the success of those practices to the truth of the ontological presuppositions that they make.

If we can justify the ontological presuppositions of area-specific scientific methods on the basis of the effectiveness of those methods in practice for achieving their goals (accurate predictions, unifying explanations, and so on), perhaps the argument can be extended to justify the ontological presuppositions of science in general, to the extent that *all* the sciences take certain foundational assumptions for granted. If at a general level, schemas of scientific explanation can be distinguished on the basis of the ontological assumptions they make, then the relative success of those explanation schemas could plausibly be taken to indicate the soundness of those assumptions. We could look to the history of scientific explanations, and base our judgments about the plausibility of various ontologies on the adequacy of the explanations of which they are a component. (PR) indicates a possible way to confirm Ernan McMullin's belief that "it would be simply inadmissible

---

[74] By a "successful" theory, I mean the usual things: that it is instrumental in the creation of technology that works, that it generates, on the whole, predictions that pan out in experiment, that it makes a part of the world more intelligible to us by unifying an apparently disparate range of phenomena; Cf., James Robert Brown, *Smoke and Mirrors: How Science Reflects Reality* (New York: Routledge, 1994), 16. A second variation on the argument goes like this:

(R*) The entities postulated by theory T exist → T is successful

T is successful

∴ The entities postulated by theory T (probably) exist.

to exclude the history of science as a major source of evidence, to suppose that we could restrict the issue of ontology to the commonalities of pre-scientific experience."[75]

The history of philosophy is littered with ontological castaways such as Leibniz's monads and Descartes' vortices—to name but two. In the end, the reason why they were dismissed as illegitimate pretenders without real existence was that the kinds of scientific explanations it was possible to formulate using them were quite limited. The history of scientific practice may not be an ideal data source for ontology, but we are likely to get a lot further with it than we would simply spending a few nights in front of the fireplace. That being said, such strategies yield only inductive support for the ontological commitments being assessed. This limitation is not restricted to Cartwright's arguments, but to any argument that takes as its aim the demonstration of the epistemic utility of the concepts, without giving a direct argument for their referentiality. If certainty is desired for an ontological claim, it will have to come as a result of a direct argument for the referentiality of the concept in question.

The implicit ontological commitments of science may be to natures, but it is still possible that science could be understood without them. If Cartwright thought that the argument established beyond a reasonable doubt that natures exist, then there would be no need for her to contrast her view to a Humean one as she does,[76] with the aim of showing why a Humean account of science is inferior. The existence of reconstructions of science with more austere ontologies (such as van Fraassen's "constructive empiricism") remains a logical possibility. While methodological arguments might cumulatively lend strong inductive support for a more robust ontology than the Humean would allow, at best it shows that the Humean standpoint is implausible. The requirements of refutation are stronger. A demonstration of outright falsehood would be necessary, and that is not what methodological arguments can give us. It may be that inductive arguments are the best we can get; it may be that we are rationally warranted in rejecting a Humean reconstruction of science in favour of a realist one.

But I think that the foregoing shows that when methodological arguments, which establish epistemic utility, are combined with metaphysical arguments that establish referentiality, together they establish that an ontological commitment to powers and natures is a presupposition of *reason in general*. An ontology of natures and their powers is, in the final analysis, the *only* one that can prove its philosophical mettle.

---

[75] Ernan McMullin, "Capacities and Natures: An Exercise in Ontology," *Boston Studies in the Philosophy of Science* VIII(1971): 49-62.

[76] Cartwright, "Aristotelian Natures," 56-60

CHAPTER FIVE: CAUSALITY, NECESSITY AND DETERMINISM

Causal realists have generally maintained that a relation of necessitation holds between causes and effects; given that some cause was present, the only effect that could have happened, under the circumstances, was the one that did. This is a *metaphysical* doctrine—there is something *real* in the world in virtue of which the effect is made to happen. There are basically two different theoretical positions available here. One holds that the necessity of causal relations is the logical necessity that attaches to statements of causal laws. In this case, the causal relation is an intensional relation between real universals, and the necessity is transmitted "downwards" to the extensional contexts, *i.e.,* from the universal to the particular. The other option holds that the necessity of particular causal relations is inherent in the "fabric" of nature itself. Causal necessity is a reflection of the ontological *determinacy* of concrete particulars, not the *universality* of causal laws. The former approach is realistic in a Platonic sense only. Aspects of the latter "Aristotelian-empiricist" approach have been endorsed by a well-known recent account advanced primarily by Rom Harré and E. H. Madden.[1] In this chapter, the Aristotelian variant is scrutinized and found to be inconsistent or ambiguous on several matters. I present a reformulation of the thesis of causal necessity based on the essentials of their approach and which, I contend, removes the difficulties of their presentation.

Causal necessity theories have often been accused of at least an implicit general commitment to determinism—that is, that nature is ultimately deterministic through and through. While that may be true of some causal necessity theories, it is not true of the version of causal necessity that I advance below. One question which looms large is how causal necessity could be present in a world which also contains chaos, quanta, and volition. Below, I sketch little more than an outline of how a realist theory of causality ought to address these issues. It is just enough, I believe, to vindicate the realist's ontological project and demonstrate its compatibility with the "troublesome" facts listed.

There are two basic types of argument presented by causal realists for the reality of causal necessity, just as there were two basic types of argument presented regarding the validity of the concepts of powers and natures. There are ontological arguments on the one hand, and methodological arguments on the other. In the most general terms, the ontological arguments attempt to secure a sense of necessity by consideration of general conceptual considerations, from which it is shown that the denial of causal necessity entails a contradiction. The methodological argument for causal necessity attempts to show that the reality of causal necessity is a presupposition of the rationality of induction. Using John Foster's recent attempt to present a methodological argument for causal necessity as an archetype of this style of argument, I show how, assuming that laws are conditional statements, these kinds of arguments fail.

5.1 Causal Necessity: The Ontological Argument

Harré and Madden's theory of causal necessity is an extension of their theory of powers and natures, and is part of what I call their "ontological replacement" project. Conceptualizing the natural world in terms of an ontology of powers and natures is just as good as its Humean empiricist counterpart at explaining the "logic" of science, while it avoids its paradoxes and philosophical problems (arising from what I argue is the sensationalist-phenomenalist current running through much anti-realist thought.)

---

[1] Their colleagues, Peter Hare and James Humber deserve note here. The work of this "school" shows the influence of earlier writings by H. W. B. Joseph, Curt J. Ducasse, and Sterling Lamprecht.

§1. *Types of Necessity*

Harré and Madden's theory of causal necessity is presented against the background of a broader conception of modality that includes "transcendental," "logical," "conceptual" and "natural" necessity. Causal necessity is treated as a species of natural necessity, and expresses the idea that the consequences of the action of a powerful particular (whatever they are), acting in accordance with its nature, are necessitated. According to the authors, there is a single sense that attaches to the concept of necessity in this context—that the concept is univocal.

> In the system we are constructing we recognize four modes of necessity, two conceptual, and two natural. We believe that the concept of necessity is univocal, that its sense is always the same but that the contexts of and grounds for its application are very various. In each major context there are appropriate grounds for attributions of necessity. We recognize the differentiated grounds and univocal sense of the concept by speaking of 'modes' of necessity.[2]

The sense of necessity which is common to the different modes of necessity is that which refers to there being *no possible alternative to the actual*. The "actual" abstracts from the particulars of fact and refers to actual outcomes, actual truth-values, *etc.*[3] The first distinction the authors draw is between modes of necessity attributable *a priori vs.* those attributable *a posteriori*.

There are two *a priori* modes of necessity: the "transcendental" and the "logical." Transcendental necessity refers to those aspects of the world which are such that if they were any different, the world would be unintelligible. The conditions for the very possibility of knowledge of the external world are those that are necessary in the transcendental sense. In the final analysis, this sense of necessity is of virtually no significance to the theory, while logical necessity plays a crucial role as a foil to the concept of natural necessity that Harré and Madden develop. Logical necessity is that property of a proposition in virtue of which the truth-value of the proposition is determined by its syntax. Truth-values can be determined either by the logical form of the proposition itself, as in the case of formal tautologies or self-contradictions, or in conjunction with other propositions, in the case of a conclusion of a deductively valid argument.

Likewise, the authors posit two modes of *a posteriori* necessity, including "natural" and "conceptual" necessity. Conceptual necessities reflect *de re* modalities; the "natural necessity in the world is reflected in a conceptual necessity in discourse about the world."[4] There are a variety of necessity claims that are then made, including ascriptions of natures to substances, ascriptions of capacities to substances, ascriptions of essential properties to substances, statements of property co-instantiation, and statements of causal laws. These are all conceivably cases of natural necessity, but it is not obvious how these natural necessities might be brought to bear on clarifying the notion of specifically *causal* necessity.

§2. *Defining Causal Necessity*

Before we can assess the soundness of Harré and Madden's case for causal necessity, we need to say with sufficient precision, what causal necessity is. In the authors' words, "causal necessity is to be understood as the modality embodied in the concept of what *must happen* when powerful

---

[2] Harré and Madden, *Causal Powers*, 19.

[3] Ibid., 19. Similarly, possibility would refer to their being some possible alternative to the actual.

[4] Ibid., 6.

particulars act in accordance with their natures."[5] This formulation is very general—almost to the point of ambiguity. A more perspicuous formulation might be:

> **Causal necessity** is the modality associated with what, when a powerful particular acts in accordance with its (specific) nature, *must happen as a consequence* of its having that nature.

This way of phrasing the claim puts the focus on the relationship between the nature of a thing and the role of that nature in the determination of what must happen—but *to what*? To it, or to other things? An entity's nature can be regarded as the cause of its own actions or (possibly) as the cause of the actions of *other* entities. Similarly, an entity's actions can be regarded as the effect of its own nature, or the effect of the actions of *other* entities. To the extent that an entity is metaphysically active, its own capacities will contribute to what it does on any given occasion. From this perspective, the entity itself (in virtue of its internal structure and processes) can be regarded as a *cause*—either as a contributing cause of its *own* action, or as a contributing cause to the actions of *other* entities which it affects.[6]

These considerations imply that *what must happen as a consequence* of an entity's having the nature that it does can refer to:

1. The necessity associated with the *performance* of action. In this sense, causal necessity pertains to the modality associated with what, when a powerful particular acts in accordance with its nature, *must happen to it as a consequence of its having the specific nature that it does.* This interpretation regards causality as the expression in action of the intrinsic structure and processes of the entity performing the action. Such an expression is necessary, because the intrinsic nature of a thing is necessarily what it is.
2. It may refer to the necessity associated with the *production* of action. In this sense, causal necessity pertains to the modality associated with what, when a powerful particular acts in accordance with its nature, *must happen to other entities as a consequence of its action*. This interpretation represents causality as a productive *process*, and identifies causal necessity with the necessary time-evolution of a causal process, whereby the action of one or more entities determines the action of others.[7]

The first sense of necessity is the one involved in the authors' claim that "the relations between what a thing is and what it is capable of doing and undergoing is naturally necessary."[8] Natural necessity characterizes the relation between the nature of a particular and its causal powers. In other words, it is just those powers that supervene on an entity's persistent structural-configurational properties—the properties that the thing has "by nature"—which are necessary. Of course, the concept of a causal power identifies what it is causally *possible* for an entity to do, not what it is necessary for it to do. In order to clarify the concept of causal necessity, some discussion of the relationship between the concepts of possibility and necessity are indicated at this point.

Recall Harré and Madden's analysis of causal powers from Chapter Four, in which they maintained that "*X* has the power to *A*" means "*X* can do *A*, in the appropriate conditions, in virtue

---

[5] Harré and Madden, *Causal Powers*, 141 (italics added).

[6] To the extent that an entity is metaphysically passive, its actions are caused by the antecedent action of other entities, and its actions are *effects.*

[7] According to Harré and Madden's univocity thesis the *same* sense of necessity is involved in these two cases. As will become clearer later, this claim cannot coherently be maintained. Indeed, the difference between these two senses of necessity, and the different kinds of arguments that are called for by the need to secure the validity of the notion of causal necessity point to more insidious troubles within their account.

[8] Harré and Madden, *Causal Powers*, 14

of its intrinsic nature." (Notice that the meaning of the power ascription is not explained in terms of what *X necessarily* does, or must do.) It will be helpful to restate the analysis of the modal force of causal powers propositions in terms of possibilities, as follows:

(**M1**) Possibility

It is possible for *X*s to do *A* (in virtue of being *X*s)

Means either

(i) It is possible for *X*s to perform actions of type *A* (under some circumstances, by some means), or

(ii) It is possible for *X*s to cause events of type *A* to occur (under some circumstances, by some means).

In order to determine what it is that Harré and Madden mean by natural (*i.e.*, causal) necessity, we need to canvass their illustrations of what is *im*possible, and why. The following two pairs of equivalent claims show that we can identify necessity claims and impossibility claims:

(a) It is impossible for *X*s to do *A*.

It is necessary that *X*s not do *A*.

(b) It is impossible for *X*s to not do *A*.

It is necessary that *X*s do *A*.

The general approach that Harré and Madden take appears to confirm the assumption of symmetry between the truth-conditions of necessity and impossibility claims:

> … the assertion that a particular has a certain nature, and that the occasion for the manifestation of its power has occurred, together with the denial that the power has been manifested, is inconsistent and this reflects the natural necessity that obtains between the nature of a particular and the occasion for exercising its power, on the one hand, and the manifestation of that power, on the other.[9]

The claim here is that if the conditions for the actualization of a capacity of an entity are satisfied in its causal context, it is a *de re* impossibility that the power-actualization processes might not occur. The same holds true for liabilities as well, as Madden's "fiery furnace" example illustrates:

> If a man falls into a fiery furnace, is it necessary that he be incinerated? Of course not, since he might get out before being badly burnt, or he might have on an asbestos suit. But say he cannot get out—the door blew shut behind him—and he had on only ordinary clothes. Now is it physically necessary that he burn to death?[10]

Madden's answer is that since the man is liable to be burned to death, and there are no factors present that would compensate for his liability to be incinerated by the fires, then it is necessary that he be so incinerated. (Ducasse's example of the decapitation causing death is exactly parallel here.) Beyond whatever observed constant conjunction between incineration and subsequent death has been observed, the justification for thinking that the death is necessary is ontologically grounded. Natural substances and processes such as fires have certain properties that they possess necessarily—"a fire" would not *be* fire if it were not hot and flaming. The causal necessity claim is therefore grounded in a claim about necessary property co-instantiations which reflect ontological

[9] Ibid., 132.

[10] Ibid., 47; Madden, "Fiery Furnace," 68.

dependency relations within the natural world.

> Here is another illustration of the sense of impossibility that the authors have in mind:
>
> It is not just a matter of fact that barns don't float off their foundations, it is, in ordinary circumstances, impossible. And that impossibility is derived from the fact that the heavy barn is in still air, within a uniform and stable gravitational field.[11]

The suggestion is that while for the most part, barns are not liable to float off their foundations, there are circumstances (*e.g.* in tornadoes) where that very phenomenon is possible. The same would be true in the case of capacities: if the conditions for the actualization of a capacity are absent, or inhibitors for such actualization are present, then the action—the power-actualization—is not possible under the conditions that obtain. Impossibility claims are therefore always to be relativized to causal contexts in which certain conditions obtain. In other words, true necessity claims are always *contextualized*. Let us present the analysis of causal necessity a bit more formally.

(**M2**) Necessity

(a) It is (contextually) necessary that *X*s do *A*

≡ It is (contextually) impossible for *X*s to not do *A*

means either:

(i) It is impossible that *X*s might not perform actions of type *A* (under conditions C), or

(ii) It is impossible that *X*s might not cause events of type *A* to occur (under conditions C)

(b) It is (contextually) necessary that *X*s not do *A*

≡ It is (contextually) impossible for *X*s to do *A*

means either:

(i) It is impossible for *X*s to perform actions of type *A* (under conditions C), or

(ii) It is impossible for *X*s to cause events of type *A* to occur (under conditions C).

(**M2**) picks out a class of cases in which it is in fact possible under some conditions for *X*s to do *A*, but not others. The impossibility of the occurrence of type *A* events under conditions *C* is not due to a general failure of *X*s to possess the relevant capacity, but due to the context-dependence or context-sensitivity of an *X*'s disposition to do *A* in conjunction with a local failure of the causal context to satisfy the conditions necessary for *X*s to do A.

According to the authors, for any token entity of type *X*, there exist conditions C in which it is naturally necessary that an action of type *A* is performed or caused to happen. That is consistent with the failure of the conditions C to obtain sometimes or even for the most part. The claim is merely that it is *sometimes* the case that *A* is necessitated—that is, when an *X* exists at a time and place where conditions C are satisfied. As I have been urging, '*A*' in the most general sense refers to a causal process—a continuous transformation of the state of the token *X* (usually, though not necessarily, through its interaction with other entities.) If such a process is a naturally necessary process, it evolves in such a way that each stage in the process is necessitated by the antecedent stages in the process. An entity *must* produce the effect that it is capable of producing, "in favorable

[11] Harré and Madden, *Causal Powers*, 12.

conditions and in the absence of constraints."[12]

Harré and Madden also discuss what appears to be a second class of cases of causal necessity in which there is "necessary relationship between the nature of a particular and the powers—capacities and liabilities—which this nature helps explain."[13] There exists "a natural necessity between what a thing is and what it is capable of doing and undergoing, and it is this natural necessity that the conceptual necessity of the concept of cause reflects."[14] No references are made to the favorableness of the conditions, the presence of triggers, absence of constraints or inhibitors, *etc.* No reference to causal context is involved. If there is one sense of causal necessity that is contextual, and another sense that is not, it is hard to see how Harré and Madden can insist on the univocity of necessity—on their being a *single* ontological tie that binds.

I contend that a plausible realist theory will drop Harré and Madden's univocity claim. The first kind of necessity, that defined in (**M2**), is the necessity associated with *power actualization*. Power actualization is a temporal process—it is the mode of causality characterizing the temporal evolution of causal processes, *i.e.*, of "efficient causation." The necessity is contextual, and pertains to concrete, singular causal sequences. If a sample of frozen H_2O is heated to more than 0°C (at 101.3 kPa), its liability to melt will become actualized—necessarily. The second kind of necessity is that pertaining to *power supervenience*. The supervenience of causal powers is causal only in the sense of Aristotelian formal and material causality. It pertains to the relation between "what a thing is (its nature), and what it is *capable* of doing and undergoing (its powers)." As was discussed in Chapter Four, the supervenience of powers on natures ensures that an entity has a capacity even when it is not exercising it, so long as it retains its nature. Ascriptions of capacity are abstract relations between structural and configurational properties, and the kinds of effects that such properties make possible. As such, they hold *across* causal contexts.

5.2 Troubles with Natural Necessity

Harré and Madden do in fact recognize the distinction that I make. However, they do not consistently recognize and apply it throughout their treatment of natural necessity—and this gets them into at least two problems early on. Raymond Woller[15] has criticized this uneasy dualism in the Harré-Madden account of necessity. His critique hardly conceals his disdain for the account overall, which is accused of being "multifarious." Woller's critique is superficial[16] in some respects and hasty[17] in others. Nevertheless, on two points in particular it is highly suggestive, and points to

---

[12] Ibid., 47.

[13] Ibid.

[14] Madden, "Fiery Furnace," 67.

[15] Raymond Woller, "Harré and Madden's Multifarious Account of Natural Necessity," *Philosophy of Science* 49(1982): 616-632.

[16] Cf. Joseph Smith, "Woller on Harré and Madden's Account of Natural Necessity," *Kinesis* 13 (1984): 77-86.

[17] Woller's claim is that it is "vain to insist that apparent counter-instances to laws of nature are spurious because once a thing ceases to behave as do others of its kind, it is no longer really a thing of that kind—of that nature." (624 n10) If this methodology were applied consistently, we would supposedly (with equal alacrity) call non-white swans non-swans. This objection seems to me to be too hastily drawn. Harré and Madden give a plausible and fairly detailed account of how scientific definitions of copper and hydrogen were decided (Harré and Madden, *Causal Powers*, 21-24, 124-26) in which the identification of what properties are definitive of the natural kind adverts to empirical discoveries of which properties explain the most other properties and powers of the substances in question.

two substantial difficulties: (a) the infinite regress of explanatory levels, (b) the failure of the realistic ontology to establish the modal force of "necessity" in causal production, or performance. After identifying why these difficulties arise in Harré and Madden's account, I indicate the form that a revised realist theory of causality must take if it is to avoid such difficulties.

§1. The Explanatory Regress Objection

Harré and Madden maintain that there is a causal necessity between the nature of a particular and its powers and capacities. So long as a particular remains the same in all causally relevant respects, it can be expected to manifest the same kind of behavior over time. If a particular entity (say a sample of sodium) has a stable identity over time, its powers will persist. It is well-known, however, that as many things (including living organisms and chemical substances) age, their powers change. Over time, they may oxidize, their chemical structures may decompose into their constituent elements, they may dissipate energy, or perhaps their structural integrity may fail. Underlying such changes in powers, there will always be a change in the underlying structure or composition of the entity. That this is the case is guaranteed by the supervenience of powers on natures—there can be no change in powers without a change in nature.

Chemical decomposition, aging, and other phenomena of transformation are what they call "non-fundamental" changes. If an entity undergoes some such transformation, we can still explain it within a theoretical framework whose ontology includes more basic elements whose nature remains invariant through the transformation. Following Madden's example, call the set of basic elements and the laws governing their behavior *S*. Is *S* necessary? According to Harré and Madden, it is not:

> It remains true that no matter of fact exists necessarily …. It *is* logically possible that fire turn into flickering lights and that human beings turn into stone—nothing necessarily is what it is—but it is impossible for a fire or human being to lose its systematically related characteristics and still remain the same particular it was before. There is no necessity that the actual be what it is, but, given what actually is the case, it is necessary that certain events and states occur in the appropriate circumstances.[18]

Harré and Madden would have been wise to jettison this tenet, for it gets them into considerable difficulty, as I now proceed to show.

If the entities described by theory *S* are themselves liable to change, then if such changes themselves are to be explained, there must be yet some *more* basic level of structure in terms of which we can account for the changes in *S*. The ontology of the world is conceived according to the hierarchical "layered" model, where powers supervene on more basic structures, and the powers of those underlying structures supervene on yet more fundamental structures, and so on. By hypothesis, the necessity that a thing invariably reproduce the same effect upon the recurrence of the appropriate conditions is contingent upon the stability of a thing's properties. The stability-in-being or the persistence of the explanatory structure is what makes it such that only one thing could happen. But the stability in the properties of a thing, given this layered model, is the stability of emergent properties, which is contingent upon the stability of the basal properties, or the structures upon which the emergent properties depend. These properties may be emergent themselves, and so on—with the hierarchy of levels of structure continuing to infinity.

It may well be a contingent fact about our world that there is no utterly stable "ultimate" level of brute facts—no level of permanent, structureless, indivisible, unchanging "Parmenidean"

---

[18] Harré and Madden, *Causal Powers*, 47; Madden, "Fiery Furnace," 68.

individuals.[19] Yet Harré and Madden's theory of causal explanation and their basic assumptions imply it. The reasons are these: if change is to be completely explained, it is to be explained in terms of the factors that make its occurrence necessary. But that which is liable to change is ontologically non-basic, and cannot figure in a complete explanation. (It is an "Aristotelian individual.") Natural necessity is therefore dependent upon the existence of Parmenidean individuals whose powers are non-emergent, indeed whose properties and powers are identical. As the authors maintain:

> References to entities which are Parmenidean halt the regress of explanation, provided that we take the final step that makes them ultimate, that is, close the logical gap by identifying their powers and their natures. The regress is halted for two reasons: (i) having reached that which is unchanging in nature, and hence in powers, there are not changes to be explained, (ii) since the natures of ultimate entities are their powers, no further characterization of such particulars is possible, because there is no independent question as to their natures.[20]

If it is correct to say that causality involves necessitation, the logic of the Harré-Madden assumptions implies that we need to *know* that there are Parmenidean individuals which can secure the ontological grounds for necessity claims. Yet, the authors admit that whether any given set of entities taken as "fundamental" really are so is a contingent, empirical matter. (Pinning the existence of causal necessity on a *contingency* of scientific discovery[21] might be construed by some as a *reductio ad absurdum* of the position.)

While this epistemic difficulty is worrisome, the crux of the explanatory regress problem is the dilemma that it poses. The dilemma is this: either there are no Parmenidean ultimates, or there are. If there *are no* Parmenidean ultimates, the robustly ontological sense of necessity for which the authors seek grounds cannot be had, and thus no causal necessity claims for singular causings can be justified. If there *are* ultimates, and their behavior is stable and necessitated, it does not necessarily follow that non-ultimate composites' behavior is. As Woller[22] notes, to "transfer necessity from the ultimate parts of things to the whole is to commit the fallacy of composition."

It is difficult to see how to assess the prospects of causal necessity from this perspective. The explanatory regress is clearly something that it would be desirable not to have to deal with. Can it be avoided? As I suggested at the outset, the root of the trouble for Harré and Madden is their assertion that a "change in the course of nature" is possible. The motivation for this position could not simply be the Humean one that such a change is "conceivable." The grounds for such possibility, given everything else they maintain, would have to be ontological.

In order to see why they take this route, we need to isolate the key problematic assumption they make. According to Harré and Madden,

> What particulars are liable to undergo and what they are able to do are determined by their natures since they are manifestations of their natures—and hence to talk about particulars remaining the same and yet lacking their usual capacities and powers is at once to assert and deny that a certain object or sample of material has a given nature.[23]

Such talk would indeed normally be taken as incoherent. Now, suppose that in some concrete

---

[19] Harré and Madden, *Causal Powers*, 161ff.

[20] Ibid., 162.

[21] Presumably this would consist in actually finding the fundamental particles—particles that would not break apart into "smaller" bits no matter how close to the speed of light we accelerate them.

[22] Woller, "Harré and Madden's Multifarious Account," 626.

[23] Harré and Madden, *Causal Powers*, 13.

circumstance in which some capacity, whose actualization we would under similar circumstances anticipate, fails to occur. Suppose, for example, that we find that a sample of copper begins to resist bending, or a sample of salt dissolved in water fails to liberate hydrogen. Harré and Madden maintain that in these cases, the sample of copper is not copper anymore,[24] the salt is not salt anymore. The necessity of the substances behaving the way they characteristically do is assumed to be *definitional* in some cases. In such cases, we do not ascribe the exception to the failure of the context to satisfy the conditions for the actualization of the power, but to the failure of the identity of the thing to remain stable over time. Here is an example of the necessity they attach to the manifestation of essential properties:

> If this stuff is as a matter of fact hydrogen, and if hydrogen is as a matter of fact an aggregate of atoms each of the same structure, and if that structure accounts for both the valency and for the spectrum of that element, then if it has this set of chemical properties it is not possible that it does not have this spectrum and vice versa.[25]

If the obtained emission spectrum of the gas sample (assumed to be hydrogen) does not match the characteristic spectrum of hydrogen, we are to assign the discrepancy *not* to a failure of the causal context to satisfy the conditions necessary for the actualization of the power, according to Harré and Madden. Rather, we have grounds for concluding that either (a) we were in error in our assumption that the gas was hydrogen, or (b) some sort of transmutation has occurred and the gas which formerly was hydrogen has changed into a form not capable of emitting the typical H^+ spectrum. If option (a) is ruled out, then (b) is the case, and we must account for the transmutation hypothesis. If we can't account for the transmutation in terms of our best present theories, then it must be the case that our best present theories are "incomplete" in the sense that there are aspects of reality that are not identified by the terms, laws, operators, *etc.* defined in our current theories.

The assumption that Harré and Madden seem to make is that if a power of a particular supervenes on an exemplification of a property essential to things of its type, or the power itself is included in the set of attributes of things of that type regarded as essential (and therefore definitive), then that power *cannot be blocked*. In other words, powers that supervene on essential properties are context-invariant with regard to their actualization. Failure of capacity-actualization is *ipso facto* proof of transmutation. Here is a clue that this assumption is in play:

> The continuity in existence of a member of a kind, or a sample of a material substance, is referred to its continuing to have the properties that constitute the real essence of the kind or material. It is notorious that the nominal essence of a material substance, say a chemical element, may be manifested only in special conditions.[26]

The implication is that the *real* essence is manifested in all conditions in some form or other. But this assumption is false in general. Essentiality of properties, and context-invariance of power-actualizations are two quite different aspects of things. We are not entitled to expect that because a property of a substance is definitive, that the capacities to which it gives rise will have special conditions of actualization. To illustrate: whatever the real essence of light is, we assume that the power of light to travel at velocity *c* is based on something intrinsic to what it is to *be* light—an essential property. But does that imply that light never fails to travel at velocity *c*? The velocity of light is *c* in a vacuum, but may be less than that, depending on the optical density of the medium

[24] Ibid., 14.

[25] Ibid., 131.

[26] Ibid., 102.

through which it is travelling. When light slows down and refracts, we do not think that the light has changed to something essentially different, and then lament the inability of our current theories to account for the transmutation.

Harré and Madden begin with the axiom that an entity must act in accordance with its nature. But an entity must also act in accordance with the conditions in which it exists. Actions of entities are the *joint* causal product of (a) the actions of other entities, (b) the nature of the entity acting, and (c) the circumambient conditions in which the action occurs. The explanatory regress problem is a result of neglecting the context-sensitivity of entities' actions, so that the possibility of a failure to act in the expected manner would be ascribed to radical alterations in the nature of things—which then need some deeper level of explanation. By affirming the consistent, general context-sensitivity for causal production, there is no need to suppose that changes in the course of nature are "possible."

There is a different necessity attaching to abstract power ascriptions than characterizes the necessity with which concrete effects are produced by causes. To think that the *same* necessity which characterizes the simple connection between an entity's nature and powers also characterizes the complex connection between cause and effect obscures the context-dependency of power-actualization, and leads to mistaken claims such as this:

> It is just the same ontological tie that binds together sequential events and states as binds together the co-existing properties and powers of things and materials into naturally necessary clusters.[27]

This is baffling not because it is metaphorical (we are all well acquainted with Hume's "secret connexions") but because it seems so plainly false. The necessity attaching to capacity-ascription statements and the necessity of statements of singular causation derive from quite different relations in reality, as I have argued. The former class of statements refers to the connection between natures and powers, while the latter affirms the connection between concrete actions or events. The natural principles of co-existence and the natural principles of succession constitute two distinct modes of causality: the first, formal and material; the second, efficient.

Harré and Madden fail to consistently separate, on the one hand, the "necessary relationship between the nature of a particular and the powers, capacities, and liabilities which this nature helps explain"[28] and on the other, the natural necessity by which generative mechanisms "bind the whole world process"[29] into a necessary stream of events.

§2. *Generative Mechanisms as Ontological 'Ties'*

In §5.2, I indicated the types of necessity included in Harré and Madden's account of natural necessity, and attempted to give a precise characterization of the logic of attributions of necessity—what they do and do not entail. We are still no closer to showing whether any attributions of causal necessity are justified. In order to show that an event was causally necessitated, it would be necessary to show how, given a description of an entity's state, its powers, and its causal context, only one action would have been possible as a consequence. In other words, we will need to show that there is some necessitating cause *X* for which the following is true:

---

[27] Ibid., 131.

[28] Ibid., 47.

[29] Ibid., 131.

> Entity *X*, when in state S, necessarily causes action A to occur (by means of P), in any causal context in which enabling conditions C_E for the actualization of the capacity are present, and inhibiting conditions C_I are absent.

Suppose we define a necessitating cause as an entity *X* for which the following is, instead, true:

> Entity *X* when in state S, necessarily causes action A to occur (by means of P), under any conditions whatsoever.

The kind of necessity that Harré and Madden are proposing does not entail this latter deterministic thesis. On their account, causal necessity is a real phenomenon even if, contingently, there are no necessitating causes in the second sense. Of course, their account is consistent with this sort of necessitating cause, but does not require their existence in order to sustain the weaker, former claim.

Harré and Madden's account contributes to the case for causal realism by presenting an analysis of the causal "by means of P" relation. For Harré and Madden, 'P' is a "generative mechanism."

> Between the nature of a thing and what it can do and undergo there is an ontological tie that binds, and this tie is to be understood in terms of the generative mechanism that produces the specific response of the particular in given circumstances.[30]

The generative mechanism is the final piece of metaphysics needed by their version of causal realism, as the following passage declares:

> It is through the application and development of the concepts of power, nature and generative mechanism that we show that a variety of rational constraints upon logical possibility can be constructed so as to limit our expectations as to what patterns of events are likely to be identified ….[31]

Knowledge of generative mechanisms by which events are produced, as cited in scientific explanations of regularities, is knowledge which serves as an objective constraint on hypothesis-formation, and limits the rational range of what events we anticipate could result as the outcome of some process. According to Harré and Madden, we can show that an event was necessary by characterizing it in such a way as to reduce "the field of the possible [to] one.'"[32] Any reduction in the possibilities is a step toward necessity—a "surrogate"[33] for natural necessity.

In this subsection, I provide an elaboration of what a generative mechanism is, and in §5.3.3, attempt to show how the concept is deployed in the method of "possibility reduction" in order to validate the ascription of necessity to causal relations. Whereas powers and natures are the concepts that establish the natural basis of possibility, "the generative mechanism is the basis of natural necessity."[34] The concept of generative mechanism first appears in Harré's discussion of causal laws. One example of a generative mechanism is the etiology of disease produced by an invasive organism. If the source of infection, and the life and reproductive processes of the infectious agent can be tracked, and if the effect of the organism upon the body at the cellular and biochemical level can be observed, then we have direct observation of the generative mechanism of the disease. The causal process by which the powers of the organism to cause disease are actualized is the genera-

---

[30] Ibid., 53.

[31] Ibid., 6.

[32] Ibid., 130.

[33] Ibid., 131, just what a "surrogate" is in this context is quite opaque.

[34] Harré, Principles of Scientific Thinking, 113.

tive mechanism of the disease.

Another example of a generative mechanism is the process by which the detection of sounds is achieved. When Hume[35] maintains that he has no idea of the connection between a vibrating string and his subsequent hearing of a sound, Harré takes that to mean that Hume regarded the causal connection between these events as irremediably inscrutable. According to Harré, this is an unwarranted elevation of scientific ignorance into an epistemic principle. We know, for example, that sound waves are propagated by compression and rarefaction of air molecules, and how the delicate anatomy of the ear acts as a transducer, converting the patterns of differential air pressure into an electrical signal. We know that there are neural pathways from the auditory nerve up to the auditory cortex, and that the delivery of electrical stimulation to the cortex induces a consciousness-like experience of sound qualia. Harré insists that we ought to regard such knowledge as knowledge of the very causal mechanisms that make the productions of their effects (in this case, hearing sounds) necessary, and provide the "missing link" that Hume sought in vain among the impressions of sensation, where of course, it could not be found.

Ever more sophisticated analytic methods in science, combined with advances in laboratory equipment technology, have provided tools for probing the fine structure of the world down to sub-nanometer resolutions. Such methods and technologies have given us the tools to discover the generative mechanisms operating in nature, and provide microcausal explanations in terms of processes unobservable to the unaided senses. The integration of the fine structure analysis of causal mechanisms with how such mechanisms appear at the level of unaided observation has become a basic paradigm of scientific explanation, at least among researchers, if not philosophers, according to Harré. If we can directly intervene into, and experimentally probe the process by which causes produce their effects, there is no "secret" to the connection between them.

Harré has suggested, in effect, that the necessity with which macro-causal processes unfold the way that they do is inherited from the causal necessity associated with the way in which individual micro-causal effects are produced. Chains of cause-effect necessitation at the micro-causal level *constitute* macro-causal processes that could not have occurred any differently than they did, under the given circumstances.

Consider the sodium-water reaction again. Harré (and Madden) maintain that if sodium is placed in water, then it is impossible that hydrogen might not be a reaction product. The fact that a Humean may regard it as conceivable in the imagination that some sample of sodium somewhere might not release hydrogen from the water is not a defeater of the judgement that *necessarily*, hydrogen is a reaction product. The Humean believes that the only constraint on what is possible, as a matter of fact, is that provided by what eludes the synthetic activity of the imagination. On the contrary, to maintain that if a sample of sodium metal is put into water at room temperature, it is "possible" that the resulting reaction will produce either hydrogen, *or* the Requiem Mass[36] presupposes an invalid concept of possibility. Of course, either it is impossible that hydrogen might not be a reaction product of sodium meeting water, or it *is* possible. The best evidence that it is possible would be a single counterexample in which a sodium-water interaction produced no hydrogen.[37]

---

[35] Hume, *Enquiry*, 51.

[36] Harré and Madden "Natural Powers and Powerful Natures," *Philosophy* 48 (1973): 210.

[37] In his critique of *Causal Powers*, Raymond Woller (Woller, "Harré and Madden's Multifarious Account," 620) does not give any evidence for the latter disjunct; instead he focuses on showing that Harré and Madden have not shown the former disjunct is true.

§3. Concretization

If individual actions are necessitated, as the realists want to maintain, it would be a serious blow to the plausibility of the position if there were no worked-out procedure for how in principle we go about identifying *what* it is that reduces the possibilities (to one). What is the logic of the reasoning by which we reach a conclusion that such-and-such *necessarily* occurs? What information is relevant to justifying claims of the form "under circumstances C, action A by entity *X* in state S, was necessary?" The classical causal realists maintain that there is a logic to the justification of sole-possibility claims, understood in terms of "reducing possibilities." The most plausible construal of their claims for necessity, I contend, casts them as claims about what can be achieved through the process of "concretiztion."

Concretization is the inverse of the process of the process of abstraction. In abstraction, specific quantitative determinations are rendered indeterminate on purpose, by reconceiving them as variables (properties) which must have some value, but may have any (within a range.) In concretization, the values of the variables are specified, so the determinacy is reintroduced.[38] Concretization can be viewed as a two-stage process.

The first stage establishes the possible actions of entities irrespective of their particular circumstances by justifying their power-ascriptions. Yet this first stage is only a "possibility-reduction" stage in the loose sense of eliminating Humean pseudo-possibilities by rejecting the broad Humean "conceivability" notion in favour of the narrower realist one. At this stage, we are warranted in holding claims such as "*anopheles* mosquito bites can cause malaria." The warrant for holding claims at this stage is no different than the warrant for holding any scientifically derived causal claim regarding the capacities and liabilities of entities of a kind—of a specific nature. The causal claims involved are power ascriptions—attributions of the capacities and liabilities of things of a certain nature. This abstract relation between a thing's nature and its supervenient powers is the necessary "conceptual connection between the concept of its nature and the specification of its reactive properties."[39] (To the extent that our knowledge of the powers of things is theoretical rather than empirical, there will be residual uncertainty about the truth of the claims in question, or about the proper characterization of the unobserved aspects of things that give rise to the powers that they have. In that sense, empirical knowledge is more effective at possibility-reduction.)

The second stage of possibility reduction involves the contextualization or concretization of the capacity-ascription. If in a given situation, all the facts that are relevant are known about the entities and processes in that situation, then we know that what happened under those circumstances could not have gone any other way. To know that they *could* have gone another way, we would have to know that we were ignorant of the presence or magnitude of some factor that had a bearing on the outcome. If we are justified in eliminating the possibility of ignorance in a specific case, then the outcome that happened was necessary.

I will continue with the example of the mosquito-causing-malaria example, since this illustrates Harré and Madden's dictum quite well, that is, that "the relationship between … successive

---

[38] Woller questions the coherence of the notion: "Such talk is at best infelicitous since what is eliminated in a reduction seems to be both impossible (because it *is* eliminated) and possible (because it was included in the set of possibilities being reduced). … The talk of reducing possibilities is misleading. It de-emphasizes the appeal to an additional sense of 'possibility,' and it thus encourages equivocal thinking about necessity." (Woller, "Harré and Madden's Multifarious Account," 618). This objection is meaningless, since what is eliminated in a reduction is *possible*.

[39] Harré and Madden, *Causal Powers*, 45.

events or states is naturally necessary when they are understood by scientists to be related in fact by generative mechanisms, whose structure and components constitute the essential nature of the permanent things and materials in the world."[40]

Characterizing the etiology of diseases is perhaps paradigmatic of identifying generative mechanisms, and of the process of possibility-reduction. At each stage in a causal process there is some physical or chemical interaction or reaction that is linked to a subsequent stage. We start off with the claim that "anopheles mosquito bites can cause malaria" which is equivalent to "anopheles mosquito bites have the power (capacity) to cause malaria." This capacity claim might be associated with some statistical claim, such as "x% of persons bitten by an anopheles mosquito contract malaria." It is not necessarily the case that if a person is bitten that they will contract malaria, so in order to say that in a given case it is *necessary* that malaria symptoms will be manifest, we need to identify the mechanism by which the bite causes the symptoms in a particular case. The stages of malaria etiology are roughly as follows:

1. The mosquito punctures the skin with an appendage designed to inject anticoagulant and extract blood from the avian or primate donor.
2. A number of protozoan parasites of the genus *Plasmodia* are transferred to the human donor from the mosquito.
3. The *plasmodia* enter the bloodstream in sufficient numbers to begin attacking red blood cells.
4. Hemoglobin is extracted from the red blood cells by the parasites.
5. Red blood cells in sufficient numbers lose the capacity to carry oxygen or die.
6. Bodily tissues fail to receive sufficient oxygen or blood vessels are blocked by accumulated dead cells.
7. Fever, chills, sweating and in rare cases coma and death follow.

It is possible (that is, it *sometimes* happens) that (1) could happen without (2) happening, in which case, the rest of the process would not be completed. It is possible that (1), (2) could occur, but in insufficient numbers, *etc.* However, in the sense that is physically impossible the (6) could occur and (7) not occur, (7) is necessary, given (6). Once we have identified the generative mechanism, we can see that the effect was necessary given the operation of the mechanism by which a causal process evolves. The "ontological tie that binds"[41] sequential events together, is thereby revealed.

The process of concretization removes the abstractness inherent in capacity claims by specifying all the relevant details of the concrete causal process by which an effect is produced and the relevant details of the causal context for the effect to occur. Given the necessity claim

> entity *X*, when in state S, necessarily causes action A to occur (by means of P), in any causal context in which enabling conditions C_E for the actualization of the capacity are present, and inhibiting conditions C_I are absent,

the concretization specifies the generative mechanism (the nature of 'P'), the state S of *X*, and whether the conditions C_E and C_I are satisfied. If all the conditions are satisfied, then *X* necessarily causes A to occur. By contextualizing the capacity ascription, we can determine, in principle, whether A is necessary.

---

[40] Ibid., 130.

[41] Ibid., 131.

§4. Does Concretization Secure Necessity?

The "*in principle*" qualification is a necessary hedge here, because clearly there will be, for many if not most situations, a lack of the relevant information or knowledge that would be *in fact* required to justify the necessity claims in question. Often what we end up with statistical information that, with some skill, we can work up into probability claims. Moreover, that is the best we can manage most of the time. In the biological sciences and social sciences, where the number of variables and processes that intersect at any one juncture in the causal nexus is epistemologically unmanageable, probabilistic claims will typically be the most precise claims we can get. Since we will, in such cases, lack the relevant information, we will not as a matter of fact be able to reduce the number of possibilities to one, which Harré and Madden maintain is required for the complete justification of a necessity claim.

Establishing necessity claims presupposes a completed process of *concretization*—the specification of the full identity of an entity and its causal context at a specific time. When we specify these aspects of the entity's causal context by introducing more determinacy into the characterization of its circumambient conditions, the range of possibilities we can rationally project as being available to an entity narrows in proportion to the degree of determinacy we specify in the context.

Even if concretization captures the logic of whether we can confirm a causal hypothesis in a specific case, it seems to fail to get at the concept of single-case necessities for the following reason. Suppose we grant that there is a necessary connection between natures and powers. For example, there is a necessary connection between being dynamite and having the power to explode. But this does not mean that every stick of dynamite will explode when lit. Wet dynamite won't explode, for example. Therefore, it is not *necessary* that dynamite explode. Whether it does or not is contingent upon a variety of contextual factors. Yet Harré and Madden insist that if a stick of dynamite does explode, that the event was necessary; and if a stick of dynamite does *not* explode, that event was necessary, too! Objection: if every event is necessary, then the concept is empirically empty—unless some events are contingent, the concept of "necessity" designates no distinction in reality. Furthermore, it is exhausted by the concept "actual." Anscombe made a similar point by emphasizing that the sense of necessitation is best understood in terms of "pre-determination:"

> It might be said that anything was determined once it had happened. There is now no possibility open: it *has* taken place! It was in this sense that Aristotle said that past and present were necessary. But this does not concern us: what interests us is *pre*-determination.[42]

If an event only becomes necessary when it becomes actual, then there is no sense in which we can justifiably say that the event was determined *before* it happened. If the necessary is in general identical to the actual, then the actual does not imply the necessary (or the modal, more broadly). If an event *A* happened at *t* is there any time $t^* = t - dt$ at which *A* was the *only* one that could have happened? If *A* was the only one that could have happened, that is assuming that the causal context at t^* was identical in every respect to the causal context at t, and that the causal process involving *X*'s production of *A* was allowed to evolve toward its inexorable conclusion of *A* without the intervention of anything to hasten or delay it. If we know that nothing could possibility affect the outcome during the interval from *t* to t^*, then it seems that the necessity claim can be sustained.

There is nothing about causal contexts in general, however, that guarantees the required stability. I would contend that such contexts are probably relatively rare. How our understanding of

---

[42] Anscombe, "Causality and Determination," 97.

causal necessity stands given this fact is nicely illustrated in terms of Anscombe's "Galton Board" example:

> Consider a steel ball dropping between two pins on a Galton board to hit the pin center under the gap between them. That it should balance on this pin is not to be expected. It has two possibilities; to go to the right or to the left. If you have a system which forces this on you, you can say: 'There has to be a determining factor; otherwise, like Buridan's ass, the ball must balance.' But if you have not, then you should say that the ball may be undetermined until it does move to the right or the left. Here the ball had only two significant possibilities and was perhaps unpredetermined between them. … We see that to give content to the idea of something's being determined, we have to have a set of possibilities, which something narrows down to one—before the event.[43]

This example illustrates a case in which the possibilities are only reduced to one at the time the action occurs, and not before. I interpret this example as showing that there is no description of the situation before the ball-pin interaction such that there is a law whereby we could predict whether the ball will go to one side or the other. The path that each such ball takes will only be decided in the interval of a few nanoseconds during the pin-ball interaction. This does not imply, however, that the path of the ball is not the only one that, under the circumstances, was possible.

If this rejoinder fends off one objection, the contention that concretization secures necessity is liable to another objection. We agree that, at the macro level, the connection between individual events may well be unknown. The objection denies that shifting the focus to the micro level from the macro level establishes necessity. If we cannot ascertain relations of causal necessity at the macro level, simply reducing the scale to the microscopic will not by itself alter the modality of whatever causal relations are discovered. This objection presumes that it is the shift from the macro to the micro level that introduces a change in modality, and this is, of course, not the case. It is the shift from the abstract to the concrete that introduces the determinacy, and hence the necessity associated with causality.

If full concretization gives us full determinacy, partial concretization will give us just partial determinacy. The range of possibilities may be reduced in such cases, but not to one. In contexts where we can identify a range of possible outcomes and a distribution over the outcomes such that some are in the long-run (under the same circumstances) more or less likely than others, then we can assign probabilities to each of the specific possible outcomes, conditional on the state of the circumstances of the entity whose actions are under investigation. The implication of this picture is that the concept of probability is epistemological. Contexts which give rise to judgements about probabilities are contexts in which incomplete information is present. When we say, therefore, that we are only rationally entitled to hold that there is a distribution of possible outcomes, rather than a single possible outcome, we mean that we are lacking information about particular features of the entity's nature or about causally relevant features of its context.

There are only two senses of indeterminacy[44] that the causal realist needs to recognize. The first is due to the unavoidable omission of information in the characterization of the identity of an entity, or in the characterization of its causal context. On the assumption that the information we have about the identity of a specimen is accurate, and as exhaustive as the science of the day provides, it may still be incomplete in the sense that under novel conditions, novel behaviour may

---

[43] Ibid., 97-98

[44] The sense in which some event is possible or not is altogether unlike the sense in which some event is probable or not; disputes about the interpretation of probability are not relevant here.

be detected; the detection of the novel behaviour is *possible*. Similarly, unknown features of the entity's environment may affect the entity in previously unanticipated ways, and the detection of *that* novel behaviour is possible.

The second sense of indeterminacy is due to the purposeful omission of information, which is key to the processes of abstraction. An abstract representation is such that it, by design, allows for the *possibility* that it may be exemplified in any number of ways within a delimited range. This is the same sense as we speak of the possible values of a variable, defined by a function over a finite domain.

I have offered an interpretation of the meaning of causal necessity claims, and given an interpretation of Harré and Madden's notion of possibility-reduction in terms of concretization. The assumption of the methodology was that causal necessity was associated with the determinacy of the thing, its state and causal context. We saw, however, that there are circumstances where the concretization is insufficient to justify necessity claims in the single case, because of a residual lack of knowledge concerning the precise circumstances of the effect. It would therefore seem that we need a broader, more fundamental argument for the reality of causal necessity. In §5.4, I present an interpretation of the central argument as it appears in the *locus classicus* of causal realism, and in §5.5 assess the impact of the available criticisms from quantum mechanics, chaos theory and free will.

5.3 Classical Causal Realism: The Central Argument

A perfectly controlled experiment is one in which if the very same conditions are replicated and the experiment is repeated, the very same results will—*must*—be recovered. However, it is not the reproducibility of the results that is the evidence for the existence of necessity, since the actual does not imply the modal. What, then, *is* the argument for causal necessity? Harré and Madden present the ontological argument for causal necessity in these two brief passages:

> To talk about particulars remaining the same and yet lacking some of their powers and capacities is to assert and deny at once that that particular has a certain nature. It is true that a particular can do and undergo different things at different times, but it cannot do and undergo those things which are incompatible with its own nature and still remain the same particular. It is physically impossible for any substance to act or react in a way that is inexplicable given its nature.[45]

> What particulars are liable to undergo and what they are able to do are determined by their natures since they are manifestations of their natures—and hence to talk about particulars remaining the same and yet lacking their usual capacities and powers is at once to assert and deny that a certain object or sample of material has a given nature. … It is physically impossible for a substance to act or react incompatibly with its own nature. It is not impossible for an object or sample to act and react differently at one time rather than another. But in general it *cannot* do so under the same circumambient conditions and be deemed to have remained the same substance.[46]

The conclusion of the argument—the causal necessity thesis—is that something's acting contrary to its nature is *de re* impossible: "actions of powerful particulars [must be] in accordance with their natures."[47] Blanshard concurs—"it seems barely possible that we are suffering from a deep and

---

[45] Harré and Madden, *Causal Powers*, 52-3.

[46] Ibid., 14.

[47] Ibid., 67.

distorting illusion in thinking that A's acting as it does follows from its being what it is."[48] While the nature of a thing provides one set of constraints that determines a range of possibilities for action, every separate moment of a entity's existence may well be associated with a different set of "circumambient" conditions, some aspects of which may be causally relevant to the character of the consequences of an entity's actions. The conditions that actually obtain in any particular (spatio-temporal) causal context provide the conditions that constitute a second set of constraints on the possible actions of an entity. Together, the nature of the entity and its causal context jointly make *necessary* the consequences of the action of any entity, anywhere.

By emphasizing the *ceteris paribus* clauses Harré and Madden include in their statement of the thesis we obtain a more precise statement of their position:

(a) An entity must act in the very same way at two different times, *if* the causal contexts are the very same, and the entity retains its identity over time.

(b) Two numerically different but qualitatively identical entities must act the very same way, *if* their respective causal contexts are the very same.

Harré and Madden's argument for these conclusions is a *reductio ad absurdum*. Suppose we maintain the opposite of either of these conclusions. Suppose that we find that the very same entity, under the very same conditions, acts one way at one time, and a different way at another time. Or suppose qualitatively identical entities act differently under the very same conditions. In either scenario, what would have to be the case? It would have to be the case that there is no necessary connection between what an entity is, and what it does—no connection between its nature and its powers. An entity which has one set of behaviours over a certain period of time might, *ceteris paribus*, acquire an entirely new set of behaviour, completely *sui generis*. To take this possibility seriously is to believe in magic—in the possibility of events for which no account can be given.

§1. The Law of Identity

Harré and Madden's argument appears to be a linear descendant of one offered by H. W. B. Joseph in his *Introduction to Logic*. For the remainder of this chapter, I will refer to any position consistent with the relevant parts of the Joseph-Harré-Madden line as "classical causal realism." Joseph's basic idea is that the "Law of Universal Causation" is a corollary of the "Law of Identity." Joseph's formulation of the Law of Universal Causation (or simply "The Law of Causation") is a conjunction of two theses, and states that "every event has a cause, and that no change occurs except under conditions with which its occurrence is connected universally."[49] To say that every event has a cause implies that there are no uncaused events. To say that every event *qua* change in nature has some cause or other does *not* place a constraint on the possible actions of entities sufficient to imply that causes *necessitate* their effects. On the other hand, to say that no change occurs except under conditions with which its occurrence is connected universally means that whatever occurs does so under conditions given which nothing else could occur; if those conditions are repeated, then the same effect will occur.

To sort out the tangle of claims involved and their relations, I propose to adopt the following formulations:

(**C**) "Every event (or action) has a cause."

---

[48] Brand Blanshard, *Reason and Analysis* (La Salle: Open Court, 1964), 471.

[49] Joseph, *Logic*, 402.

(**C+**) "An entity must act in the very same way at two different times, *if* the causal contexts are the very same, and the entity retains its identity over time," or "if two numerically different but qualitatively identical entities must act the very same way, *if* their respective causal contexts are the very same." (Principle of Uniformity of Nature)

Short version: "Same causes, same effects."

(**C++**) **"**Every event has a nomologically sufficient cause." (Law of Causation)

Thesis **C++** is a combination of both **C** and **C+** and is equivalent to Joseph's Law of (Universal) Causation. **C+** is the Principle of Uniformity of Nature, and corresponds to my (a) and (b) above. Joseph claims that **C++** is entailed by the Law of Identity (**LI**), so we should now turn to his explanation of this, and determine its truth or falsehood.[50]

The argument depends on the necessary truth of the Law of Identity, so we first need to assess the plausibility of the case for this premise. The law states, simply, that whatever is, is what it is—"A is A." While traditionally one of the three "Laws of Thought," the Law of Identity, like the Law of Contradiction, reflects an "apprehension of a necessity in the being of things."[51] "It is because what is must be *determinately* what it is, that I must so think." **LI** is an ontological principle expressing the idea that to be is to be something determinate. "A is A" can be interpreted as a truncated form of the categorical proposition "All A is A," whose substitution instances relate entities to their identities. The 'A' of the subject position is the term designating the entity considered as an undifferentiated unity, while the 'A' of the predicate position is a term designating the totality of the determinate attributes of the entity. **LI** *identifies* the totality of the determinate attributes of the entity with the entity itself. There is numerically only one entity involved in any exemplification of **LI**. It relates, under the identity relation, two different *perspectives* on entities—one which regards entities as ontological unities, one that regards them as integrated ensembles of distinguishable characteristics. This is the meaning of saying that a thing is *what* it is.

What considerations could be offered for regarding **LI** as a necessary truth? One argument is that it is rationally inescapable. As the foundation of logic, it is the underlying formal principle of the validity of demonstrative reasoning, irrespective of subject matter. Any argument directed against the truth of the Law of Identity which laid claim to validity would be in the embarrassing position of presupposing the truth of **LI** in the argument for its falsehood. This fact renders any argument against the Law of Identity incoherent.

It is also a presupposition of reason in another sense, for according to Joseph, "if … there can be no qualitative sameness in what is numerically different … there is no reasoning."[52] Since conceptual thought by its very nature involves the identification of abstract qualitative identities among numerically distinct particulars, thinking about (not simply recollecting) particulars is implicitly general. We may regard particulars as all ontologically dissimilar, while at the same time being able to identify sets of relationships *among* the particulars that give rise to ways of regarding particulars as exemplifications of natural kinds. (This does not necessarily imply a realistic view of the ontology of universals, although realism concerning universals is consistent with it.)

Another point that can be offered in favour of the Law of Identity—it is *self-evident*. The very act of perceiving the particulars we might thereafter conceptualize presupposes that we are able to

---

[50] A variation on this theme is presented in Mario Bunge, *Causality* (New York: Meridian, 1957) in which he calls **C** the "genetic principle" and **C+** "the causal principle."

[51] Joseph, *Logic*, 13.

[52] Ibid., 409.

identify differences between them and other things in their vicinities, and similarities to other things in our past experience. Such discrimination of similarities and differences is discrimination of similarities and differences *in their properties.* If things were not intrinsically fully determinate as **LI** implies, then perceptual discrimination would not be possible—there would be nothing in virtue of which we could recognize particulars as being different from others. Clearly, there are determinate ways in which perceived entities differ from their backgrounds. The **LI** expresses the fact of this self-evident and inescapable determinacy of things. The Law of Identity has an important ontological corollary, which I call "determinatism"[53] (as opposed to "determinism"). Determinatism is the view that for any property that an existent has, it has it in some specific, particular value. It is an implication of the idea that to exist *is* to exist *as* something specific. This is the Aristotelian idea that there is *no concept of existence as such, applicable to subjects of an indeterminate identity*.[54] The Law of Identity is generally regarded to be equivalent to the Law of Excluded Middle; the truth of one entails the truth of the other and *vice versa.* Metaphysical realism entails that determinatism is true, where determinatism can most clearly be expressed in one of the following forms:

(T1) For any property Φ for which ϕ is an exemplification of Φ, if some entity X has Φ, then X either does or does not exemplify ϕ at a time t.

(T2) For any quantitative determination ϕ_o of a property ϕ, if some entity X has ϕ in any measure, then either X does or does not have ϕ_o at a time t.

(T1) states just that if something has a property, it exemplifies it in the form of a concrete instance of that property. For example, if being a collie is an exemplification of being a dog, and Banjo is a dog, then either Banjo is or is not a collie. (T2) states that if an entity has a property that comes in exemplifications that carry quantitatively specifiable values, then it has that property to some degree or another at any time. For example, if a collie can be measured in terms of how much it weighs, then either the collie weighs, for example, 20kg or does not weigh 20kg. Let us call the first principle "type determinatism," and the second "token determinatism." Here are some straightforward applications of token determinatism to some physical attributes which seem trivially obvious in normal contexts, but are contentious in the quantum domain (where it is often suggested classical logic simply does not apply):

1. For any A, existing in (Minkowski) space, its spatial boundaries are determinate. For any entity A, and any point in space (x,y,z,t) relative to some frame of reference, part of A either *does or does not* occupy that position.
2. For any entity A, and any magnitude of mass m, either A *does or does not* have mass m (at velocity v) at time t; its mass is determinate.
3. For any entity A and velocity v, A either *is or is not* travelling at v at time t; its velocity is determinate.

Determinatism as an ontological doctrine has no epistemic implications—it does not imply whether we can measure the magnitudes of the properties ascribed to A. All that it says is that if A is *real*, the

[53] The distinction here has also been characterized in a rather ugly way as the distinction between "evolutionary determinism" and "vertical determinism." It is the latter position that corresponds to "determinatism." Note that one can be a realist about properties but not be a "determinatist." It is possible to reject the view that all properties are determinate and still be a realist about properties. Such people might, for example, hold that property *universals* are indeterminate to a degree but real, in the relevant sense.

[54] Of course, Aristotle was writing before Bohr, Heisenberg, *et al.*, whose own positions contradict the universality of application of the Law of Identity. One can hardly resist supposing that Aristotle would have regarded the Copenhagen school as expounding a mathematically sophisticated version of Heracliteanism.

attributes are determinate. To deny this is to take the "real" out of "*real*ism."

The foregoing considerations ought to be persuasive in favour of the necessary truth of the Law of Identity. What then of the claim that the Law of Causation is a corollary of it?

§2. The Law of Causation

To say that every event has a cause may seem to be a simple, uncontroversial truth. If **C** is a corollary of **LI**, then to say that every event has a cause means that it is *logically* impossible for there to be uncaused events. Unfortunately, Joseph does not make clear the logic of the argument for this corollary; he seems to assume that it follows from the presuppositions of the intelligibility of reality as a whole. If so, then it is not from **LI** that **C** follows, but they both follow from antecedent considerations of the intelligibility of the world and the requirements of reason in general. Joseph argues that **C+** follows from **LI**, but to argue that **C++** *also* follows from **LI**, he would also need an independent argument for **C**, which he does not give. If there is a valid direct line of reasoning from **LI** to **C,** the argument would have to be one made out along the following lines.

Let us suppose that entity *a* underwent some genuinely uncaused event *Q*.[55] An uncaused event *Q* for an entity *a* means that in some causal context, *a* underwent some change, but was not made to change by anything. That implies that *a* changed in the absence of antecedent or simultaneous factors in its causal context upon which the change was conditional or dependent, *including its own nature*. Now, if *Q* was an actual action for *a*, *Q* is a possible action for *a*. But if *Q* is independent of *a* and independent of *a*'s causal context, then any other action *Q** would seem to be equally possible for *a*. If *any* action at all is possible for an entity, then there is nothing—no properties—specific to *a* in virtue of which *a* is determinate, or determinately what it is. If so, then *a* has no identity—it is nothing in particular. According to **LI,** such an *a* does not exist. This is just the claim that Joseph maintains: that a hypothetical entity with no identity would be nothing in particular, and therefore would not a possible candidate for real existence.

This *reductio* refutes the idea that it is logically possible for there to be uncaused actions. The logical impossibility of uncaused events follows straightforwardly. **LI** indeed does imply that there are no uncaused events (imaginative fancy notwithstanding). That, however, is only half the thesis. In addition to the explanation of why **LI** entails **C+**, we also need an explanation of why **LI** entails **C+**, if we are to secure the thesis of the nomological sufficiency of the cause of any event.

Here is Joseph's statement of the entailment relation between **LI** and **C+**:

> It is no more than a corollary of the Law of Identity, that the same thing unaltered on different occasions, or two things of the same nature, should under the same conditions produce the same effect.[56]
>
> If one thing the same in nature at different times, or two things the same in nature, are to act in situations the same in their nature, they must act on both occasions in the same way. This is not a generalization from experience: it follows from the sameness of thing and of situation.[57]

In short, *given the same cause and conditions, the same effect must occur*. It is consistent with the non-repeatability of particular states of affairs; the hypothetical character of the principle neither requires nor implies the existence of empirical regularities:

---

[55] The only candidate I've ever seen for an uncaused event in the literature is radioisotope particle emission.

[56] Joseph, *Logic*, 420.

[57] Ibid., 406.

> What we call single things are highly complex, and their properties and behaviour depend upon their composition, and upon the situation in which they are placed relatively to other things; we may believe that whenever one complex thing of precisely the same kind is placed in precisely the same situation as another, it will behave in precisely the same way; nor is more required by the Principle of the Uniformity of nature; and yet we may doubt whether such precise repetition ever occurs.[58]

In order to support the contention that **LI** implies **C+** Joseph uses the *reductio* strategy:

> to say that the same thing acting [twice]… under the same conditions may … produce a different effect, is to say that a thing need not be what it is. But this is in flat conflict with the Law of Identity. A thing, to be at all, must be something, and can only be what it is.[59]

> To suppose that the same cause—other things being equal—can have different effects on two occasions is as much as to suppose that two things can be the same, and yet so far their attributes different. … So far as conditions precisely the same in kind recur, they must, if there is such a relation as cause and effect at all, have the same effect.[60]

Given that **LI** is true, to say that every event has a nomologically sufficient cause is equivalent to saying that every event has a single, possibly non-recurring set of causal determinants which are jointly sufficient to produce the event. It is virtually certain that no entity is completely insensitive to its environment, depending entirely on its own internal processes for its actions. Any entity stands in some relations to others that interact with it in some respect and to some degree, if by nothing more than the force of gravity, for example. Nor is any entity's action completely determined by antecedent external events. For any entity, the way in which it acts must be regarded as a partial expression of what it is, and a *partial* expression of what its environment is—that is, the causally relevant parts thereof that constitute the causal context of the entity. Provided the causally relevant aspects of its environment are stable, it could only act differently if it *were* different—if the entity were not what it determinately is. *This* is what violates the Law of Identity, according to Joseph. Changes in the ways in which things act over time must, according to this view, be systematically related to changes in the character of the external conditions into which the entity enters. No variation from such a pattern—"anomalous" behaviour"—is to be accounted for by magic, miracles, or a metaphysically irreducible random element in nature. The explanation will advert to the only possible alternatives: to either a previously undetected change in the intrinsic state of the entity or in its causal context.

One objection that is likely to be raised at this point would be the following. If Joseph says that the ways in which things change over time are to be explained with reference to some systematic relation to its intrinsic states and its extrinsic conditions, then any law expressing such a systematic connection would have the same epistemic status as any other empirical generalization—it would be just as open to refutation by future counterexample. Since a proposition asserting a matter of fact is never immune from falsification, attributing the modality of necessity ("must" language) to such a proposition, *i.e.*, causal generalization, is unwarranted; it fails to acknowledge the contingency of empirical knowledge claims.

There are a couple of responses to this objection available. The first is this: Joseph is *not* claiming that causal laws are contingent empirical generalizations, or regularities. Rather, causal

---

[58] Ibid., 402.

[59] Ibid., 408.

[60] Ibid., 409.

laws express *hypothetical uniformities*. Regularity-statements, on the other hand, are statements about non-exceptionless, general patterns, for which probabilistic analyses are inescapable, because variations in contextual circumstances (the presence of inhibitors of the actualization of a capacity in some causal contexts, for example) affect virtually all entities. An actual uniformity is the "limit" of a regularity, which is only empirically obtained in the perfectly controlled experiment. Uniformity statements are hypothetical: if an entity acts in such-and-such a way in this context, then in an identical context, the entity will behave in an identical way. *If* we can reproduce the causal context of an entity's action under experimentally controlled conditions, we can elicit the uniformity as a regularity.

The causal laws that make claims about what happens for the most part, or tends to happen, or happens with a certain long-frequency, are indeed empirical generalizations, and these do not support the modality claims that causal realists make for causal laws. If the sorts of causal laws for which the causal realist claims necessity are not empirical generalizations, then just what *are* they? According to Joseph,

> Uniformity of action is not indeed the fundamental element in the causal relation, for it depends on repetition of the action; the causal relation has nothing to do with *number* of instances, so far as its *existence*—though much as far as its *detection*—is concerned; it is bound up altogether with the *nature* or *character* of things, and the nature of anything is not a question of the number of such things that may be or have been fashioned.[61]

The causal relation is supposed to obtain not among entities' natures and their characteristic actions, so that the laws which express *de re* necessities are of a different formal character than causal laws. On this point, Harré and Madden introduce a useful distinction, which helps to formulate the second response to the objection.

If, for example, the claim is made that "an acid solution turns logwood solution red,"[62] we can take this as an implicit example of a causal law. On the issue of the propositional form such laws take, we need to note that there are at least two quite different approaches available. The first, and more familiar, is to render it in the extentional idiom: "for any solution, if it is acidic, then if added to aqueous logwood, the logwood turns red." This sentence expresses the accidental co-instantiation of the predicates over an open-ended range of instances. The second approach to Harré and Madden's example is to render it in the intentional idiom, that is, to construe it as affirming a conceptual connection between being an acid and having the capacity to turn logwood red.

The intentional interpretation responds to, and incorporates the requirement that the causal relation express a connection between the nature of the cause and the nature of the effect. More specifically, the causal law "acid solution turns logwood solution red" expresses a connection between, on the one hand, the capacity of acids solutions to turn logwood solution red, and on the other, the liability of logwood solution to incur a determinate colour change.[63] It does not affirm that in *every* conceivable causal context, acid solution will turn logwood solution red. There, may,

---

[61] Ibid., 407.

[62] Cf. Harré and Madden, *Causal Powers*, 8-10.

[63] An alternative formulation of the point states that the predicate "has the capacity to turn logwood solution red" is to be regarded as a partial explication of the *meaning* of the predicate "being an acid solution." This affirms the abstract conceptual relation between the predicates does not express anything definitional about acidity, since the capacity of acids to produce changes in pH indicators and other substances derives from more causally fundamental chemical characteristics of the substance in question.

for example, be logwood solutions that are buffered to attenuate the powers of the acid to below the threshold H^+ ion concentration necessary to elicit the usual chromatic response. This would be an example of a unfavorable causal context, *i.e.*, a context in which the conditions are unfavorable to the actualization of the entity's capacities. To whatever extent the actualization of a capacity of any specific acting entity is sensitive to the presence of inhibiting conditions and dependent upon enabling conditions, an unfavorable causal context can mitigate against the entity's being able to bring about the change that it otherwise would. But the fact it otherwise *would* presupposes that the causal power is active, but there are contextual factors that constrain or counteract the full realization of its action-potential. To the extent that conditions are favorable, a causal process will occur in which the action-potential of the entity will be more or less realized. If we assume the favorable causal contexts are the usual ones, then it follows that *usually*, its causal powers will be actualized.

This seems to be the sense in which Harré and Madden maintain that "the concept of causality is such that the rational response to the failure of the usual conditions to produce their expected outcome" would be to make one of the following conclusions: either (a) the instance was aberrant—some unknown factor had become part of the causal context, which could be in principle discovered and removed from the context, (b) the entity in question had, in the meantime, undergone a change in its nature (*e.g.*, a transmutation) sufficient for the disappearance of its former powers and their replacement by others, (c) our original necessity claim which was thought to hold in fact is mistaken, or (d) our world was destroyed, and immediately replaced in its totality by one that looks the same for all intents and purposes, but works according to different principles.[64]

What is *not* possible is that favorable circumstances for the actualization of an entity's powers should obtain, the entity should be qualitatively the same in all causally relevant respects, and yet the effect that would (under those specific circumstances) be an expression of the entity's nature not occur.

5.4 The Retreat from Determinism

Classical causal realism, as presented in this chapter, maintains two parallel theses: that an entity must act in the very same way at two different times, if the causal contexts are the very same, and the entity retains its identity over time; that two numerically different but qualitatively identical entities must act the very same way, if their respective causal contexts are the very same. We saw that this is an elaboration of the claim that given the same cause and conditions, the same effect must occur. Unfortunately, this position is open to a variety of objections, since the doctrine of causal necessity articulated by classical causal realism is equivalent to universal causal determinism.[65] Recall Joseph's claim that **C+** followed from the Law of Identity and the meaning of the notion of "sameness of thing and of situation."[66] If determinism follows from the Law of Identity and the meaning of the term "same," then *determinism is analytically true.*

What is determinism? Earman characterizes determinism as the doctrine that the same causes necessarily produce the same effect. It is perhaps more often presented as the thesis that all actions (and other events) are fully necessitated by (temporally) antecedent conditions. The two formula-

---

[64] Harré and Madden (*Causal Powers*, 73ff) seem serious about this last possibility, which is, unlike the others, completely arbitrary and contradicts the empiricist conception of possibility that they appear to endorse elsewhere. Cf. Ibid., 46-47.

[65] John Earman, "Determinism in the Physical Sciences," in *Introduction to the Philosophy of Science,* ed. Merrillee H. Salmon, *et al.* (Englewood Cliffs, NJ: Prentice Hall, 1992), 237.

[66] Joseph, *Logic*, 406.

tions are equivalent, as can be shown by considering any causally closed physical system.[67] If every event is necessitated by antecedent conditions, then every state of the system is necessitated in accordance with some exceptionless, deterministic law or laws by the previous states of the system. Similarly, the entire state of the system at any time will move, in "obedience" to some deterministic law or laws, to the only "permitted" possible future outcome at each point in time.[68] These facts about determinism seem perfectly compatible with Joseph's insistence that "there is not a particle of water whose path is not absolutely determined by the forces acting on it" in accordance with an "unbroken reign of law."[69] There seem to be, however, a couple of possible variations consistent with these formulations of determinism if we attempt to express the idea with a bit more formal precision. We could characterize it in a couple of different ways, each of which catches some part of the intuition behind determinism.

System S is a deterministic system iff:

(D1) For any state λ of entity *X* in *S* whose initial value is $\lambda_o=\lambda(t_o)$, the state $\lambda(t)$ evolves (nomologically) such that at any later time *t'* it has only one possible value $\lambda_1(t')$.

(D2) For any state λ of entity *X* in *S* whose initial value is of type *M*, where *X* is *M* at *t* iff λ falls in the interval (λ_-,λ_+), the state $\lambda(t)$ evolves (nomologically) such that at any later time *t'*, it has only one possible type of value *N*, where *X* is *N* at *t'* iff λ falls in the interval (λ_-',λ_+').

For the sake of convenience, let us call the first *token determinism*, and the second *type determinism*. Notice, however that type determinism is equivalent to indeterminism under certain conditions—when a narrow taxonomy of properties is chosen, such that small (*M*-type identical) differences in $\lambda(t)$ may yet produce large (*N*-type distinct) differences in $\lambda(t')$. The *deterministic limit* is the point at which the type-individuation criteria for the nomologically related properties *M* and *N* is narrow enough so that the difference between indeterministic and deterministic systems cannot be empirically distinguished.

There exist serious challenges to determinism stemming from its incompatibility with (i) the chaotic time-evolution of complex systems, (ii) with quantum mechanics, and with (iii) human volition. These challenges, combined with the question of what the "same cause, same effect" principle actually means, show that the analyticity of determinism is false.

Much speculation has attended the question of why causality is linked conceptually to determinism. Surely the intelligibility of the cosmos conferred by the success of Newtonian astronomy and its mechanical world-picture played a part. The cosmos of the 17th century was unified, ordered, simple and understandable in terms of a few basic laws. The mathematics of classical mechanics, empirically adequate as it was, applied to corporeal bodies whose dynamics were modeled by linear equations, which were simple and basically passive. They were "self-determining" only insofar as they had inertia, and manifested activity only in the ability to exert an inverse-square attraction. It is perhaps natural to account for the explanatory and predictive success of Newtonian physics in terms of its literal truth—of its getting the ontology of the world right in its theories. It would be a mistake, however, to assume that we are somehow epistemologically entitled

[67] Intuitively, causal closure is the idea that nothing supernatural can intervene to predetermine by fiat some outcome that is not determined by any antecedent state of the system.

[68] As Earman suggests, this claim in hard to make out in the context of special relativity, where there is no absolute time. There is absolute time in classical physics, but there is also the problem of breakdown of causal closure of the physical by particles escaping to spatial infinity.

[69] Joseph, *Logic*, 402.

to read ontological significance into the mathematical formalism of an empirically adequate theory. That form of inference to the best explanation has proven resistant to validation.

The attractiveness of determinism is grounded in assumptions about the nomologicality of nature, the linearity of causal processes, the relative passivity and simplicity of bodies, the predictive success of computationally tractable mathematical models of natural phenomena. Yet, its domain of applicability must be curtailed in the face of causal feedback, interaction, chaotic time-evolution, and teleological causation in living organisms. Can a plausible form of causal realism still be defended? What interpretation of **C+** can be sustained, if any? Can we square realism with the facts while maintaining its metaphysical commitment to the Law of Identity and its implication that an entity must act in accordance with its nature? If affirmative answers are to be forthcoming, there are some initial general claims that have to be endorsed at the outset, which will indicate the direction in which realism needs to be reformed in order to shed the implicit general determinism of classical causal realism.

First, entities are profitably regarded as falling somewhere on a spectrum of passivity whose ends are represented on one end by the spherical inelastic billiard ball, and on the other by the rational agent. Entities for which the "same cause, same effect" principle is roughly true *in practice* will tend to fall near the passive end of the spectrum. Second, and closely related to the first, is the matter of the ability of entities to store and transform energy for the purpose of replication. Those that can do so exhibit novel teleological modes of action (*e.g.*, biological processes) that cannot be accounted for solely by mechanical principles. Third, realists must be open to the possibility that no true, general version of the "same cause, same effect" principle can be defended on logical or conceptual grounds. It will likely turn out to be the case that the extent to which an entity of a given kind behaves deterministically or indeterministically is a function of the persistent and enduring aspects of the entity's composition and structure, and is an empirical matter. Entities that satisfy the determinists' preferences for mathematically tractable, passive, law-governed motion should not be regarded as ontologically privileged. Similarly, the electrons, persons, and thunderstorms should not be ontologically marginalized.

In the next three subsections, I sketch the nature of the difficulty traditional causal realism faces, show how these difficulties arise, and assess what modifications of realism are necessary to achieve coherence with the facts.

§1. The Chaos Challenge

There is an ambiguity in the Harré-Madden-Joseph approach that makes The Principle of Uniformity of Nature **C+** difficult to support. Joseph, for example, maintains the following:

> To assert a causal connection between *a* and *x* implies that *a* acts as it does because it is what it is; because, in fact, it is *a*. So long therefore as it is *a*, it must act thus; and to assert that it may act otherwise on a subsequent occasion is to assert that what is *a* is something else than the *a* which it is declared to be.[70]

This is a strongly-worded argument for causal determinism or universal causal necessitation; its conclusion is a blanket metaphysical prohibition of indeterminism. But the universality, or the generality of the argument's conclusion raises suspicions. What are *a* and *x*, after all?

The intended interpretation can be inferred from the chemical examples classical causal realists typically use which seem to fit this pattern perfectly. Sodium, hydrogen, copper, gold (even

[70] Joseph, *Logic*, 407.

aspirins), all have well-defined essential properties, whose failure to be manifest under the conditions in which they are defined, would support the inference Joseph makes. For example, to affirm the liability of gold to dissolve in *aqua regia*, and at the same time to maintain that *this* sample of gold will not dissolve in this bit of *aqua regia* is just the kind of incoherence that the classical causal realist thinks is true of singular causal claims in general. This argument trades on the same context-invariance assumption for power-actualizations associated with essential properties of natural kinds as Harré and Madden made, and which I discussed and rejected above. The assumption may well hold for some *a* and some *x*, but it does not contain therein a general truth about causation.

The classical causal realists can (logically speaking) be right about gold, or natural kinds with well-defined essential properties more generally, but wrong about causation as such. (And I think they *are* right about tokens of natural kinds with well-defined essential properties). What is persuasive about the argument is its combination of an abstract inference trading on a subtle ambiguity in the word "same," combined with a set of illustrations from chemistry the analysis of which are uniquely strong in their conformity to the pattern that the realist would like to have as an inductive springboard to a general conclusion about causality. The ambiguity is this: does the "same cause" mean "quantitatively identical in all causally relevant respects?" Or should "same cause" should be interpreted more broadly to mean "similar cause?" In other words, is it type determinism or token determinism that is being defended?

Let us first take the case where some entity *X* has the capacity to produce a change *A* in another entity, and it is in virtue of *X*'s being of a certain *type* that it has the capacities that it does—because it has the right structural/compositional properties. But properties are *abstracted* features of a category of particulars which vary quantitatively from one another. *X* is a member of a family of entities *all* of which have the same capacity, because they share the same structural/compositional attributes. A natural interpretation of **C+** is that the same *type* of cause produce the same *type* of effect. It is also natural to think of the particulars subsumed by a type as being *similar* with respect to the properties in virtue of which they are categorized together.

Imagine a situation where on one occasion, entity *X* produces change *A*, and yet another entity *Y* of the very same type as *X* fails to produce change *A*, while in the identical conditions. Is such a situation conceivable? According to classical causal realism, it could only be so if *X* and *Y* were in fact different in a causally significant respect. If such a situation were observed, the rational response would be to *recategorize X* and *Y* on the grounds that some causally significant property that is had by *X* is not had by *Y*, or is had by *X* to a significantly greater or lesser degree. We originally thought that the *X* and *Y* were the same; now we revise our assessment—they were merely similar. The point is that differences in the causes which we might regard as merely quantitative across a number of contexts might in *some* contexts constitute differences that result in qualitatively different effects.

Breakdowns of **C+** in which minor differences in the states of the particulars acting (under identical conditions) lead to divergent effects can always be explained away by isolating the respect in which the causes differ, and reclassifying the causes as being of *different* kinds in light of the narrowed taxonomy of causally significant properties. Whether the reclassifications required in accordance with **C+** would become *ad hoc*, and distinguish no interesting differences in properties is unlikely, since it is interesting differences in *effects* that lead us to look more closely into the nature of things. If similar causes have dissimilar effects, then it is because the entities involved have differing capacities, and thus differing properties which give rise to them. In other words, breakdowns in type determinism can always be "fixed" by narrowing the range of variation over which

the description "same type" applies.

We have been looking at the case where, assuming precisely identical conditions, *prima facie* type-identical entities of some domain *D* might behave in different ways. The breakdown in **C+** is averted if we narrow our taxonomy of causally significant properties of entities in a way that allows us to explain the divergence in effects in terms of unexpectedly small differences in the causes. Determinism is still plausible for the entities in *D* so long as there are good epistemic grounds for the increasingly subtle differentiations that are made. A similar result is forthcoming in those situations in which given identical (or near-identical) entities and actions, widely divergent effects are observed to occur. Determinism is still plausible so long as the divergent effects can be ascribed to slight differences in the conditions in which the entities are acting, and the sensitive dependence of the entities' behaviors on very small differences in the conditions in place from time to time.

There will always be regions of the world in which the systems we study display a high degree of sensitivity to initial conditions. These regions are, in fact, easily found. Minute changes in the stress across a tectonic fault may in general have no effect, but within a certain range, similar minute changes can result in a massive slippage, and the result is an earthquake. If we have a lack of information about the unstable regions of the phase space of a system or substance, then the behaviour of the system or substance may appear random or "chaotic." The more chaotic a system is, the more of its phase space is unstable, and the more "random" or haphazard it seems.

Nonetheless, it is fully consistent with *token* determinism that such chaotic behavior be present throughout the world. Take, as an other example, the airplane wing upon which we depend to get us safely to and from philosophy conferences. The strain induced on the wing in response to the stress of atmospheric turbulence is a linear function of stress, in accordance with Hooke's Law. Deformations (bending and twisting) of the wing will tend to disappear when the stress is eliminated; the wing returns to its equilibrium state. This is only true up to the "elastic limit" of the wing, however. When the stress approaches the elastic limit, the strain function becomes non-linear; higher-order terms suddenly take on large values, and the wing breaks.

Systems that are characterized as "chaotic" are those that exhibit non-linearity over large regions of their phase spaces, but are fairly rare. More condition-specific phases of instability are ubiquitous in nature, however.

We have seen that the "similar causes, similar effects" principle, or *type determinism*, is falsified whenever similar causes lead to qualitatively different effects. This can happen when the purportedly type-identical entities acting (the causes) in fact differ in some causally relevant respect, or when quantitatively different but near-identical conditions give rise to divergent effects (given identical causes) due to non-linear responses of the entities in the range of those conditions. The lesson is that **C+** should not be interpreted as "similar causes, similar effects" because whether similar causes *do* produce similar effects is highly contextual, and dependent upon empirical factors.

While **C+** is subject to empirical test—and falsification, is it nevertheless the case that a narrow interpretation of **C+** can be sustained? Is there anything implausible about (universal) token determinism, according to which "given causes and conditions that are quantitatively identical in every respect, the exact same effects necessarily occur?" One difficulty with such a principle is that it could never be *tested*, for we would have to be able to measure all physical magnitudes to arbitrarily high degrees of precision, and we cannot do that. There does not seem to be any way to derive an interpretation of **C+** that is unambiguous, true and universal. The classical causal realists were misguided in thinking that the Law of Universal Causation is analytically true, since

the narrowness of the taxonomy of causally relevant properties and hence the sense of the concept "same" in the formulation "same causes, same effects" is inescapably context-dependent. While type determinism may well be (approximately) true for some entities and some physical systems, it is not true in general, as chaotic systems clearly attest. The realist is best served by regarding causality in general as the Law of Identity applied to action—that how a thing can acts must be in accordance with its identity. This leaves open entirely the extent to which an entity's performance of the actions of which it is capable is context-sensitive, and should be interpreted as acknowledging that how a thing *reacts* must also be in accordance with its identity. Actions can be regarded as analogous to vector sums—each entity involved in the power-actualization process contributes a component to the "resultant" or the overall action, even if the process is dominated by one powerful particular.

None of the foregoing argument counts against local token determinism, however. Even those systems that exhibit high degrees of sensitivity to initial conditions, and whose time-evolution is still anomic and unpredictable, may well have a non-zero, measurable deterministic limit. We may be able to approximate token deterministic systems in highly controlled experiments, and if we can, that accounts for why we can generalize from a single instance in such cases.

§2. The Quantum Challenge

There are several aspects of phenomena at the microscopic level that are not found at the macroscopic level, and which are not capable of being assimilated to "classical" concepts of the physical world. The non-classical features exhibited by quantum phenomena are: discontinuity, complementarity, indeterminacy, and non-locality. It is the last two features of quantum mechanics which are generally considered to contradict classical notions of causality, especially universal *deterministic* causality. There are assumptions of classical or Laplacean determinism that the realist is not compelled to accept. In the last subsection, I showed, with reference to chaos, that universal causal determinism ought to be rejected, and that realism is consistent with indeterminism. In particular, the classical assumptions of the linearity of causal processes ought to be dropped, so that causal feedback, interaction and the resulting and chaotic time-evolution are countenanced within the more liberal realist ontology.

Nonetheless, the realist is still committed to the Aristotelian principle of the Law of Identity or its corollary, the doctrine of "determinatism" (see §5.4.1). In the Copenhagen interpretation of quantum mechanics, due primarily to Bohr and Heisenberg, there is a failure of actions to be determined by antecedent conditions, as well as a failure of existents to be determinate. While causal realism can countenance the first, it cannot endorse the second without incoherence. Additional experimental confirmation of the completeness of quantum mechanics and the empirical adequacy of the standard formalism is supposed to have accrued as a result of the Aspect experiments. It is widely regarded as a technical knockout of those with classical realist sympathies (such as Einstein) who, at the time, maintained that the Copenhagen interpretation was unsatisfactory.

The majority view today is that the standard formalism is adequate and that the Copenhagen interpretation saves the phenomena. My contention is that there are indications that higher-order criteria of adequacy for the interpretation of physical theories that may not be met by the Copenhagen interpretation. There are, as touched upon in Chapter 4, complications introduced by the issue of how, and under what circumstances, we can reason from the predictive success of a scientific theory to a position on the ontology presupposed by the theory. The case against a realist interpretation of quantum mechanics cannot be easily dismissed. Nevertheless, I contend that the

implications of quantum mechanics for realist theories of causality have been greatly exaggerated.

Cases where "inference to the best explanation" fails to provide any compelling reason for chosing among ontologically non-equivalent but phenomena-saving theories are cases where realistic theories acquire their greater plausibility. The "prejudices" of the classically-trained intuition hold up somewhat better under logical scutiny than is often supposed. While interpretations of quantum mechanics that are consistent with causal realism, such as David Bohm's causal interpretation, are surely no panacea, I argue that their coherence with realist presuppositions gives them a slight edge over their contemporaries. There are, presently, several different groups of theoreticians working within the Bohmian "paradigm," and nothing in my argument depends upon whether any of the various Bohm-style theories are right, only that they are not obviously and uninterestingly wrong.

There have been two main lines of attack against realist interpretations of quantum mechanics. The first was articulated in the late 1920s, and another became available in the early 1980s. We will begin with the traditional anti-realist argument.

Einstein "wanted things out there to have properties, whether or not they were measured."[71] So does the realist. Indeed, things having *real* properties is essential to being a metaphysical *real*ist. As I argued above, being a realist implies that you believe in determinatism. But determinatism is just what was called into question in the 1920s, on epistemological grounds. The strongest argument against determinatism is the argument from the Heisenberg Uncertainty Principle. We have already discussed the distinction between epistemic and ontological determinacy, where I showed that realism is attached to the idea that indeterminacy is merely epistemological. According to the Copenhagen school, the indeterminacy is not just epistemological, it is also ontological.

The standard story about how the epistemological indeterminacy arises is simply this. To predict the future trajectory of a particle, S, you need to know its position and its momentum. In order to measure the position and momentum of S, some causal process is involved. The measurement has to occur by *some* means. Usually that means that some incident energy is directed toward S, *e.g.* in the form of light or electrons which interact with S, and are picked up by a detector. The interaction between S and the measuring particles will disturb the S, so that the best measurements will only give us an estimate of S's position or momentum. The more precise the position measurement of S we want, the less we want it to scatter our measuring particles. To reduce the scattering, we need to increase their energy. But with higher energy, we disturb S's velocity vector, so our measurement of S's momentum will be less accurate. There is a trade-off here: if we want more position accuracy, we have to sacrifice momentum accuracy, and vice versa. The Uncertainty Principle states that the best we can do is to reduce the product of the measurement uncertainties to approximately Planck's constant. The uncertainty here is simply a natural consequence of the fact that to acquire knowledge of the identity of things is to do it *by some means*. Just as we cannot perceive physical objects by no means at all—we need to use our hands or eyes—we cannot measure physical objects by no means at all—we cannot do it without interacting with the subjects of measurement.

According to Heisenberg, there simply is no interpretation of the physical quantities before they are measured—all that quantum mechanics does for us is to give us probabilities that the

---

[71] David N. Mermin, "Is the Moon There When Nobody Looks? Reality and Quantum Theory," in *The Philosophy of Science*, ed. Richard Boyd, Philip Gasper and J. D. Trout (Cambridge, MA: Bradford / MIT Press, 1991), 501.

observables will have certain values. The "state vector" is an evolving set of probabilities of that which will only carry a physical meaning when the quantum state is measured. This expresses the verificationist belief in a "reality that becomes actual or definite only upon observation."[72] What is the case *before* observation is just a mathematical representation of a distribution of potentials—there is no fact of the matter about what the potentials are of before hand if there is no actual existence that they describe.

Notoriously, the justifications offered for the plausibility of a verificationist principle by the architects of the Copenhagen interpretation are difficult to fathom, when they are offered.[73] The anti-metaphysical zeal of positivism was part of the cultural milieu in the learned communities of the time, and the scientists who had absorbed their positivism by intellectual osmosis may well have not been prepared to offer explicit, sound philosophical arguments. Cushing's suggestion is a stronger one—that such arguments were (and still are not) forthcoming. The Heisenberg-Bohr interpretation of QM was in part the product of a logical error, an illicit inference from a constraint on the precision of measurements to a claim about the natural world—an inference from epistemology to ontology. This inference was "not only logically unjustified but also not demanded, either by experiment or by the formalism of quantum mechanics."[74]

For the sake of argument, let us suppose that the verificationist principle is acceptable, and that as applied to quantum mechanics, the denial of determinatism follows as a consequence. Since realism implies determinatism, if determinatism is false, then realism is also false.

There are some very odd consequences of the Copenhagen interpretation. If it is accepted that the act of measurement "collapses" the state vector, by reducing the probability wave packet to a single value for the observable, such a measurement can only be interpreted as, in a fairly literal sense, creating the very phenomena being measured. Unfortunately, this conception of measurement seems to entail a dilemma which might be taken as a *reductio* of the interpretation.

The dilemma begins: either measurement is a physical interaction between an object system and a measurement apparatus, or it is not. If it is, then the interaction should be describable in the same terms as those that apply to other interactions, as the time-evolution of the object/apparatus system ought to obey the Schrodinger equation. If so, then upon the measurement interaction, the object/apparatus system should evolve into another superposition. But this is contradicted by our experience. According to the (standard) the interpretation of the wave function, it represents the evolution of a superposition of possible (*i.e.*, non-actual, non-real) quantum states, and there is no fact of the matter about the "real" states antecedent to the measurement. How then do we explain why if an interaction is macroscopically manifest it is "allowed" to violate the description of the state vector in terms of Schrödinger's equation (since you get a specific value), but no other interactions are so permitted?

Suppose, then, that the measurement is not a physical interaction between an object system and a measurement apparatus. In that case, it must be a mental or spiritual interaction—a direct intellectual apprehension or noesis of the state of the system. This alternative is a logical possibility, and many popular writers have jumped on it as support for claims that are often ascribed to

---

[72] James T. Cushing, *Quantum Mechanics: Historical Contingency and the Copenhagen Hegemony* (Chicago: University of Chicago Press, 1994), 30.

[73] Bohr (the founder of the Copenhagen interpretation) was, as a writer, "often difficult to understand, and at time just plain opaque." (Ibid., 28)

[74] Ibid., 25.

Eastern mysticism, Taoism, or New Ageism in one form or another.

Albert Einstein opposed the emerging "received view" regarding QM. He and his colleagues in their famous EPR paper took a somewhat different anti-Copenhagen attack. They argued that quantum mechanics was incomplete, in the sense that not every relevant feature of the reality that the theory should intend to describe had a counterpart within the formalism of the theory. If QM were complete, EPR argued, it would entail that some state-descriptions for some systems would imply "spooky" non-local causal interactions. What the EPR argument was entitled to conclude in fact was something a bit weaker, namely "not that QM is incomplete, but that it is either incomplete or nonlocal."[75] Locality was a conviction that Einstein could not surrender; neither could he give up realism, especially the objective simultaneous determinacy of position and momentum of quanta. EPR attempted to use the assumption of locality to prove incompleteness. Unfortunately for Einstein, the attempt backfired.

The experimental tests of "Bell's inequalities" conducted by Alain Aspect and his colleagues in the 1980s essentially showed that the predictions of quantum mechanics were satisfied. This effectively removed the basis for Einstein's objections to quantum mechanics on the basis of its failure to secure determinacy for the physical properties of a quantum system. The Aspect experiments proved (according to most commentators) that one could not be a realist and believe in locality at the same time; realism and locality are not compossible positions.

The second anti-realist argument is rather straightforward at this point: if we are forced to choose between locality and realism, and the truth of special relativity is to be accepted without reservation, then realism is "out," and so are deBroglie-style hidden-variable theories. Of course, there is still the matter of the measurement problem to resolve (which has proven a most vexing puzzle).

There is, however, another option: the causal interpretation of David Bohm. Given Bells' inequalities, and the plausibility of realism, *non-local* hidden variable theories are still possible. Given plausible assumptions[76] "Bohm's theory recovers all of the statistical predications of standard quantum theory and has a largely classical ontology."[77] Although conventional wisdom about quantum mechanics holds that it is *in principle* impossible that position (and hence, a trajectory in space and time) is a possessed property of a microsystem, "Bohm's theory accomplishes just that."[78] In Bohm's theory, the ontology is strictly classical, in the sense that the position and momentum of particles have definite values; the "superposition conception" is rejected. Therefore, there is no state-vector collapse, and the measurement problem in its traditional form does not arise. The Bohmian interpretation of the Uncertainty Principle is strictly epistemic, probability enters the theory the same way as it does in classical thermodynamics to describe Brownian motion. Perhaps most importantly, token-deterministic chaos and non-deterministic randomness are empirically indistinguishable experimentally, since either assumption is consistent with the well-confirmed statistical predictions of quantum theory—which can be derived either from Bohm's formalism or from the standard formalism.

Despite these positive aspects of a Bohm-style realist interpretation, non-locality is a problematic feature of the theory. The Aspect experiments rule out a common-cause explanation of the

---

[75] Earman, "Determinism," 255.

[76] Cushing, Quantum Mechanics, 45.

[77] Ibid., 25

[78] Ibid., 26

EPR correlations, while a direct causal link between the separated detector-stations is ruled out by the special-relativistic prohibition on simultaneous action at a distance. It has appeared to some that the non-locality of the Bohm theory implies the propagation of causal influences at arbitrarily high velocities, which contradicts special relativity. On the other hand, the "no-signaling" theorems show that there is no way to use measurement statistics at one part of a non-locally "connected" system to get information about the settings of the detectors at another part. The non-locality, given no signaling, suggests that there is no actual propagation of *causal* influence occurring from place to place, and so no violations of special relativity. The explanation of the obtained correlations is still mysterious, however. The correlations may be the most interesting Salmon-style pseudo-processes yet discovered.

The traditional anti-realist argument from quantum mechanics is not obviously successful, because of its presupposition (at least in some of its defenders) of a verificationist premise that there is no compelling reason to accept. The Bell's inequalities and the Aspect experiment results only count against a realistic or "determinatist" version of quantum mechanics in which there are "hidden variables" if at the same time, assumptions about local causal token determinism are abandoned. It appears then, that neither the mathematical formalism or the empirical results of modern physics force us to reject realism, determinatism or the Law of Identity, as is often supposed. While there are outstanding conceptual difficulties with all of the various Bohm-inspired versions of quantum mechanics, the arguments which are used to attack realistic interpretations of quantum mechanics are themselves based on assumptions which have their own questionable foundations and troublesome implications.

Quantum realism is not a monolith (nor for that matter, is quantum anti-realism); there are various available formalisms and interpretations, and nothing in my argument depends on the details of any one version, Bohmian or not, being correct.

On empirical grounds, the realist and anti-realist interpretations of quantum mechanics have commensurable scientific balance sheets. As reflected in the Bohm and Copenhagen interpretations, they are empirically equivalent theories that employ different mathematical representations of stochastic systems to recover the same statistical predictions. Each has vexing conceptual difficulties: the former has to account for non-locality, while the latter is saddled with explaining state-vector collapse. The one asset that, in my estimation, gives realist accounts a comparative advantage is that they cohere, not with mechanistic conceptions of deterministic causality, but with high-level regulative constraints on theory-construction derived from the basic principles of conceptual consistency and intelligibility which are embodied in the principles of classical logic.

§3. The Volition Challenge

The last major difficulty for classical causal realism is its *prima facie* incompatibility with the fact of human volition, *i.e.*, with genuine freely willed conscious choice among possible alternatives. It therefore falls victim to one horn of the old freedom-determinism dilemma. More damagingly, this conclusion is explicitly argued by at least one thinker, who has argued that "freedom of the will" is an illusion, based on reasoning virtually indistinguishable from that of Joseph, Harré, *et al.*:

> Every thing-in-being must be *something*, must have a definite nature. It cannot … be like the *ens metaphysicum*, that is, a thing which simply *is* and no more than *is*, without any definitions and properties, and consequently, without a definite way of acting which flows from them. For every thing-in-being must have a nature in virtue of which it is what it is, which this being always maintains, and whose manifestations are called forth of necessity by causes …. But all this is just as true of man and his will as of all

> other beings in nature. He too has … fundamental properties what make up his character and require only an outside inducement in order to reveal themselves. Consequently, to expect that a man should act one time in one way, another time quite differently, in response to the same cause, would be no different than to expect that the same tree which bore cherries this summer should bear pears in the next.[79]

It follows from the assumptions of classical causal realism that for someone making a choice between a given set of alternatives, that person is not (qualitatively) the same person when he or she makes different choices at different times—surely an unwelcome implication! If human beings are no different in kind than any other animal product of natural evolution, then there is no way to rationally regard their mode of consciousness as an exception to the principles that govern every thing else in the natural world. There are no ontological exemptions from the Law of Identity.

According to a recently articulated position,[80] volition is "cognitive self-regulation"—an evolutionary adaptation which allows the organism possessing it to have a substantial degree of control over the function of its own consciousness. Control over voluntary motor function is causally dependent upon prior acts of volitional control over mental functioning. This is held to consist essentially of directly willed governance of the level of conceptual attention or "focus" given by the subject to his environment or to articles of prior knowledge. The choice to focus, *i.e.*, to volitionally initiate a more or less active state of awareness—is a "causal primary."[81] Metaphorically, it is the choice, when looking, how much to *see.*

The existence of a primary choice to focus which is not explained by any other choice (but which explains other consequent choices in both thought and action) is plausible, insofar as it avoids the problem that would otherwise appear: if every choice were dependent causally upon some antecedent choice, there would be an infinite regress of choices, which is impossible, given a human's finite life span. More broadly conceived, however, this position simply avoids one horn of the classical dilemma at the cost of impaling itself on the other. For if the primary choice is free, it is not necessitated by a nomologically sufficient cause, and volition violates the law of causality. The other horn of the dilemma is no more solace, for it saves causality at the price of freedom.

Suppose that right now, I have just chosen to focus my awareness on the gentle whirring sound of my computer's cooling fan. The noise has been present and available to my awareness all along, but just now, I chose to attend to its volume, timbre, and the very slight rapid modulation in the frequency of its sound. Presumably I can ask the question: was my choice caused? The same could be said for any volitionally-actuated *generalized* increase in the sensitivity of my mind to the world around me. Is any such choice caused? Either it is, or it is not.

Suppose the choice *is* caused. Now, since it is the causally primary choice (*ex hypothesi)* the cause was not something of my choosing. Hence, the cause is something I did not choose. Since causes and associated conditions necessitate their effects, my focusing my awareness was not something that I could have avoided doing at the time. In effect, the primary "choice" was necessitated—not a choice at all.

Suppose the choice is *not* caused. In this case, there is an uncaused event, which violates the Law of Causality. Since **C+** is a corollary of the Law of Identity, the latter principle's abandonment

---

[79] Arthur Schopenhauer, excerpt from *Essay on the Freedom of the Will,* in *Reason and Responsibility*, 8th ed., ed. Joel Feinberg (Belmont, CA: Wadsworth Publishing Co.), 370.

[80] Harry Binswanger, "Volition as Cognitive Self-Regulation," *Organizational Behavior and Human Decision Processes* 50 (1991): 154-78.

[81] Ibid., 168.

is the ultimate consequence of the possibility of acausal volition. The contradictions which ensue can be grasped in short order. We cannot be free to violate the laws of logic; we cannot have a nature such that we *can* fail to act in accordance with it.

In short, the choice either has a necessitating cause, in which case it was not chosen, or it has no cause, in which case it is inexplicable. A comprehensive realist theory of causality provides the resources to frame a general dissolution of the classical dilemma by exposing the unwarranted assumptions upon which it implicitly depends.

The reasons why the dilemma of freedom *vs.* causal necessity arose in the first place are two: (i) it is assumed that only events could be causal relata, to the exclusion of entities, and (ii) it is assumed that the event must be causally antecedent to the event to be explained. The first assumption was rejected on the grounds that events have no causal efficacy—that only acting agents do. If it somehow doesn't make sense to ask whether the primary choice is caused, then it is an even more stark exception to the law of causality, because it does make sense to ask about any other event whether (and by what) it was caused.

A more promising line of approach would be the agent causation approach. This is the view that a free choice is not caused by any other event, but it is caused by the agent. In this view, events can be caused not only (or perhaps not at all) by other events, but also by entities, and one might go on here to define an 'action' as an event that is caused by an entity. This enables one to maintain the law of causality as "every event has a cause" without leading to an infinite regress, since the cause is not always another event. We might then say that an action is free when it is caused by the agent, and not free when it is caused by something external to the agent.

If the law of causality says only that every event has a cause, and the cause may be an entity, then it says nothing more than that every event is the action of some entity. In that case, it does not rule out the possibility that an entity might have two or more courses of action available to it at some time—which is why this view allows the possibility of free will. But at the same time, this formulation of the law of causality also allows the possibility of chance events, such as are contemplated in most interpretations of quantum mechanics. The radioactive atom is capable of decaying, or not decaying. Whichever it does will be the action of the atom, so the law of causality is not violated (whatever happens has a cause: *viz*., the atom itself is the cause). So I think that it is essential to the law of causality that not only is every action caused by an entity, but the causal factors present (the entity's characteristics, plus its circumstances) are sufficient to determine which action it performs—something like that is what is required to rule out random events.

5.5 Causal Necessity: Methodological Arguments

It is often supposed that if causal necessity were not a feature of the world, induction would lose its warrant. Thus, the law of causation is sometimes regarded as a presupposition of the logical standing of induction. Let us see whether the problem of induction is successfully addressed by this principle.

§1. The Problem of Induction

Humean skepticism regarding the warrant of inductively established, or *a posteriori* knowledge claims, is based on assumptions whose grounds have already been challenged in Chapter One. On the other hand it would be misleading to say that a non-Humean or realist counter-analysis of causality "solves" the problem of induction because, as Harré and Madden suggest, "the

rendition of the problem is essentially Humean in the first place."[82] On reflection, however, there is no single problem of induction; the philosophical issues are manifold and intertwined. A helpful approach to formulating a tractable version of the problem was contributed by Evan Fales,[83] who distinguishes four different epistemological questions about induction that are neutral with respect to the possible theoretical positions one could take on the answers:

1. How do we isolate and identify significant patterns of regular recurrence?
2. How do we isolate all those factors that are causally relevant to what will happen in a given circumstance?
3. How do we justify the prediction of a single future occurrence?
4. How do we justify generalizations and law-statements based on finite evidence?

A couple of comments are called for with respect to this division of the issues. First, Questions 1 and 2 are essentially questions of scientific method, adequate responses to which will need to be localized to specific fields of inquiry, since they will presuppose a detailed theoretical knowledge of a specific science, and knowledge of its standard experimental practices and procedures. The most plausible general claims about inductive confirmation of causal hypotheses that can be made are claims for the reasonableness of the employment of methods such as Mill's methods of induction or linear causal modeling. What we might call the "little problem of induction" is essentially the question of how we can be sure, assuming such methods in principle work, that we have done everything right, given the fallibility and limits of our natural faculties.

The answer to Fales' Question #3 will depend on an adequate answer to #4, which itself is the "big problem of induction." Evan Fales' own attempt at neutrality is strained here, since it is the Humean formulation of the problem of induction that makes the finitude of the evidence epistemically relevant. Humean approaches to induction generally accept what Karl Popper called the "primacy of repetitions"—that the number of repetitions of a sequence of events which conforms to a generalization is relevant to belief (rationally justified or otherwise) in the generalization. According to Hume, repeated positive-instance confirmation makes a generalization naturally more appealing, even if it does not confer any additional logical probability upon the generalization. According to Rudolph Carnap, it does. According to subjective Bayesian accounts of induction, positive instances of a generalization have normative import insofar as they require a reassessment of one's posterior probability assignment to the generalization in order to maintain one's diachronic doxastic coherence. The common denominator of these versions of inductivism is that induction is inherently an *ampliative* form of inference—that induction somehow goes *beyond* experience. This intuition is caricatured by the inference "this raven is black, this raven is black … thus, all ravens are black."

Induction can go beyond experience in one of either two ways, as it is commonly understood. First, induction may ascend to generalizations which assert that *all members* of a certain class currently have a certain property, when the evidence is only that some members do. Second, induction may ascend to the conclusion that a generalization will continue to hold true in the future, when the evidence is only that it has held in the past. The "big problem of the induction"

---

[82] Harré and Madden, *Causal Powers*, 71.

[83] Evan Fales, "Causation and Induction," in *Causation and Causal Theories* Midwest Studies in Philosophy, vol. IX, ed. Peter A. French, Theodore E. Uehling, Jr., and Howard K. Wettstein (Minneapolis: University of Minnesota Press, 1984), 113-34.

arises from the clash between two *prima facie* incompatible truths: (i) that induction is ampliative in the second sense, and (ii) that the temporal stability of our generalizations is not something that we can count on. The dubious stability of our causal generalizations is just a special case of the overall problem.

The universality of the conclusion of the inductive generalization is supposed to go beyond all possible experience, and that is inherently problematic from an empiricist point of view. The only universal propositions that could be known to be necessarily true are purely formal ones, whose truth depends on formal conventions—not upon the apprehension of some relationship to reality.

§2. The Nomological-Explanatory Response to the Problem of Induction

Madden and Joseph both believe that the law of causality is a presupposition of induction; without a necessary connection between cause and effect, induction would have no internal warrant. John Foster's[84] recent "solution" to the problem of induction (*i.e.,* Question 4) falls into the category of methodological arguments for causal necessity, and is one of the more sophisticated versions of this general pattern of argument. It will well be worth a look. Foster thinks that the best solution to the problem of induction exploits the concept of natural necessity, which implies that induction presupposes necessity. He exploits the concept of "objective natural necessity" to drive the "nomological-explanatory solution" (**NES**) to Question 4, framed as an answer to the question of how accidental generalizations are to be distinguished from generalizations of law. While Humeans and non-Humeans alike think that this distinction can be made out, Humean for the most part subscribe to the formula that Law = Accidental generalization + *X*.[85]

Foster begins with an assessment of a now-familiar argument. Suppose all observed As have been Bs. Believing that the next to-be-observed A will also be B would be justified by the knowledge that all As are Bs, if such knowledge were possible. If the conclusion were justified, then nothing would be gained by replacing "all As are Bs" with "all As must be Bs," since that substitution would not strengthen the inference. Moreover, there is no evidence that would confirm "all As must be Bs" that would *not* confirm "all As are Bs," and so the additional modal strength supposedly introduced by the "must" is illusory. The plausibility of the anti-necessitarian argument depends upon whether we ought to accept the assumption that induction actually ever works this way, even schematically. Foster replies in the negative.

He is right to point out that propositions of the form "all As must be Bs" would have more explanatory power than ones of the form "all As are Bs" if the former could be justified. But since any process of reasoning from empirical data to the validation of a scientific hypothesis depends on prior theoretical background assumptions, no purely inductive path from data to hypothesis-confirmation is ever available. If there is no such "pure inductivist" approach to justifying scientific claims in general, how is the even more difficult task of justifying necessity-claims to be discharged? Foster suggests that the answer is to be found in abduction: if the stronger modal claim "all As must be Bs" is the best explanation for the data, then we are justified in accepting it.

If we consider the fact of gravitational interaction, we find that all bodies "behave gravitationally." More precisely, that means: for any two bodies A and B, A has the capacity to attract B in

---

[84] John Foster, "Induction, Explanation and Natural Necessity," *Proceedings of the Aristotelian Society* 83 (1984): 87-101.

[85] Fred Dretske, "Laws of Nature," *Philosophy of Science* 44 (1977): 248-68.

virtue of its mass, and A has the liability to be attracted to B in virtue of its mass. In cases where the capacity of one mass to attract another is not actualized, because of the presence of an inhibiting object, once the inhibitor is removed, the usual gravitational behavior is manifested. As far as we know, all bodies interact gravitationally—without exception, despite questions about the exact physical nature of the gravitational field. What is the proper interpretation of this fact? Is this an empirical regularity the expression of which carries no more modal force than any other accidental generalization? If that is so, then the regularity is an inexplicable brute fact about the world. It is not, however, a brute fact about the *nature* of the world. The fact that matter has behaved gravitationally in the past is no guarantee that it will continue to do so. No matter how confident we might be that nature will continue to behave in accordance with the observed historical pattern, it is still possible that (either locally or cosmically) a change in the course of nature will occur.

It does seem bizarre to think that this unbroken pattern of behavior—this exceptionless regularity on a universal scale—is merely a cosmic accident. As Foster observes, "surely it must be that gravitational behavior is the product of natural necessity: bodies have hitherto always behaved gravitationally because it is a law of nature that bodies behave in that way."[86] But in what sense is it true that bodies always have behaved gravitationally? Does it mean that all bodies at all times behaved in accordance with the law of gravity? If Nancy Cartwright is right, then at a sufficiently precise level of description, *no* bodies actually manifest a behavior strictly in accordance with the law of gravity. The actual situation involves two separate facts: (i) gravitational forces always contribute to the behavior of any massive object, in the sense that the gravitational force always contributes a vector component to the *overall* motion of any body, (ii) gravitational forces, in the vast majority of actual causal contexts, swamp the contribution of other forces, so that the overall motion approximates that which would be produced by the gravitational interaction alone. The first fact reflects the necessary connection between the mass of objects and their causal powers. The second fact reflects facts about the distribution of available causal contexts in which gravitational forces are more powerful than others.

We can combine these two facts by saying that all bodies *tend* to act gravitationally. A tendency-ascription combines a capacity-ascription with an assessment of how strong the causal powers are, on average. To take another example, we can say that a certain parasitic protozoan has the capacity to cause *pneumocystis carinii* pneumonia, but that it does not *tend* to do so, because it is only in specific causal contexts—*i.e.*, persons with depressed immune systems—where its capacities can be actualized. On the other hand, if the claim is relativized to those causal contexts in which the inhibitors of the capacity are usually absent, then the particular will tend to have that capacity actualized. In those contexts, the parasite tends to cause disease. It should be clear that tendency-statements do not have the modal force we expect of laws of nature—only capacity-statements, which identify necessary connections between natures and powers, have the explanatory power that being a law requires. With the proviso that by a "law" of nature, Foster ought to be thinking "capacity-ascription," let us proceed to his main argument, which appears to be this:

F1 In order to justify believing in a prediction based on an observed regularity, we need to be able to justify believing that the regularity will continue into the future.

F2 The truth of an empirical generalization based on past experience is compatible with its future falsehood.

[86] Foster, "Induction," 89.

F3 Therefore, empirical generalizations cannot provide rational grounds for belief in the persistence of a regularity.

F4 Rational grounds for belief in the persistence of a regularity is provided by the identification of factors which explain why the regularity in question has obtained in the past.

F5 The best explanation for the persistence of a regularity in the past is the presence of some naturally necessary phenomenon the description of which will involve propositions having the status of laws of nature.

F6 Such natural necessities, in virtue of being necessary, are *permanent* aspects of nature, and will by that fact, be present in the future.

F7 Therefore, only a law of nature, describing some necessary permanent feature of the structure of world, can help explain the regularity, and thus provide a rational basis for believing that the regularity will persist.

Regarding the law of gravity as the description of a naturally necessary phenomenon "eliminates what would otherwise be an astonishing coincidence: it enables us to avoid the incredible hypothesis that the past consistency of gravitational behaviour, over such a vast range of bodies, occasions and circumstances, is merely accidental."[87]

This "nomological-explanatory solution" to the problem of induction hinges on the truth of the foreseeably controversial claim that we can identify the natural laws which bear the natural necessity. Even if the truth of some postulated law would, in virtue of its modal force, constitute the best explanation for the observed regularities, it is another question altogether how we would go about discovering them. The gravity example is perhaps too easy, because the way that things behave in virtue of their mass is an observable regularity, and also exhausts the empirical content of the law. This fact about this particular example does not suggest any path to a solution for more difficult cases in which the appropriate characterization of the explanatory law involves the specification of theoretical quantities, which cannot be directly ascertained from observational regularities, but only inferred therefrom.

Consider Foster's account of the relation between the law of gravity construed as an accidental generalization *vs.* a law of nature, reflecting a natural necessity in the workings of the world. Suppose we have a universal generalization, expressing the regularity, *all As are such that under conditions N, they have behaved in manner B* (*e.g.*, all massive bodies under every condition have behaved in a manner compliant with the gravitational interaction). In order to have the power to explain the regularity, we simply recast it using modal language: *necessarily, any A under conditions N behaves in manner B.* The regularity is therefore instantly transformed into a law of nature, with the explanatory power to account for the regularity, and the rationality of belief in its future persistence. The author does express some sensitivity to the availability of alternative theoretical explanations for the observed regularity, and it will be instructive to show how his rejection of these alternatives is only superficially satisfying.

Let us grant that the fact of universal gravitation is to be subsumed under a law of gravitation if the fact is to be explained. What laws would be consistent with the observed facts? It seems that there are at least a few:

L1 It is a law that for all times, under all conditions, bodies behave gravitationally.

L2 It is a law that for all times before time *t*, under all conditions, bodies behave gravitationally.

[87] Ibid., 91-2.

L2* L2 & there is no more comprehensive gravitational law.

L3 It is a law that for all times, under conditions N, bodies behave gravitationally.

L3* L3 & there is no more comprehensive gravitational law.

The explanatory law to be preferred is the one that is the least puzzling, according to **NES**. Now, L1 is simpler than either L2 or L3, because both L2 and L3 are restricted in scope. For L2, the "fact" that there is some special time *t* (whether it is the present, or some fixed time) at which either the law goes out of effect, or there is some more general law that "takes over" surely calls for further explanation. Likewise L3 leaves the question of how, in non-N contexts, bodies act. Similarly, L2* and L3* treat the restrictions of L2 and L3 respectively as not themselves open to theoretical explanation at a higher order of unification, which adds another layer of puzzlement—if either L2* or L3* were true, would the observed regularity continue to hold? If so, then the regularity would be just a brute fact, which would leave us with the same uneasy sense that an undiscovered explanation is still lurking. If not, then we still need an explanation for what is different about non-N and post-*t* contexts. Just as cosmic coincidences are inherently suspicious because it is so overwhelmingly likely that they can be explained, restricted laws are "*a priori* perplexing."[88]

Despite the satisfaction we get from explanations which leave no loose ends, the satisfaction is superficial, and temporary, once we try running the **NES** solution with other examples, such as those in which the postulated natural necessities which account for the regularities are highly theoretical. In such cases, we cannot simply modalize the accidental generalization, even if doing so gives us the best explanation, because the terms of the law at the level of deeper explanation do not relate the same quantities as appear in the law of gravity case.

The most promising account of laws of nature regards them not modalized accidental generalizations, but as capacity-ascriptions. The explanation for why As have behaved in manner B in the past is that As have the capacity to behave in manner B in virtue of their having some intrinsic property P and the conditions suitable for the actualization of the As capacities, N, have been present in all hitherto observed contexts in which As have acted.[89]

---

[88] Ibid., 97.

[89] James Woodward's recent accounts of laws based on invariance of power-actualizations seem to be to be the right sort of realist theory. E.g., James Woodward, "Realism About Laws," *Erkenntnis* 36 (March 1992): 181-218; idem, "Capacities and Invariance."